**The Dissertation Committee for An Li Certifies that this is the approved version of the following dissertation:**

# Kinetic Studies of HIV-1 Reverse Transcriptase Nucleotide Selectivity, Drug Resistance and RNase H Activity

**Committee:**

Kenneth Johnson, Supervisor

Ilya Finkelstein

Adrian Keatinge-Clay

Rick Russell

Christian Whitman

# Kinetic Studies of HIV-1 Reverse Transcriptase Nucleotide Selectivity, Drug Resistance and RNase H Activity

**by**

**An Li, B.S., M.S.**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

**May 2016**

# Dedication

To Zhigang and Xinguo

## Acknowledgements

There are so many individuals who have made this dissertation possible and who have made my graduate career a precious memory. First, I would like to thank my loving parents, Zhigang and Xinguo, who have guided me to build the dream of being a scientist in biological sciences. My deepest gratitude is to my Ph.D. advisor, Dr. Kenneth Johnson who gave me the opportunity to explore science freely, and the guidance to solve difficult problems when my steps faltered. I thank him for molding me as a scientist who can understand, explore and communicate science with dedication, confidence and meticulousness.

I would like to thank our lab administrator Dr. JoAnn Johnson who has spent tremendous efforts to make our lab a professional and lovely place to work. I would also like to thank Dr. JoAnn Johnson for her exceptional eagerness and patience to help me sharpen the oral presentation skills, which will bring benefit to my whole scientific career.

I would like to acknowledge my doctoral committee members: Dr. Ilya Finkelstein, Dr. Adrian Keatinge-Clay, Dr. Rick Russell, and Dr. Christian Whitman for the friendly guidance, thought-provoking suggestions and scientific expertise.

Many thanks to Dr. Matthew Kellinger who laid the foundation for the projects I have worked on. Thanks also to Dr. John Brandis, Dr. Dipa Batabayal, Dr. Jessica Mckenzie, Dr. Yufeng Qian and Virginia Nguyen for their help, guidance and support during my graduate study.

I am grateful to Shanzhong Gong who has grown with me in this lab for his friendship, insightful comments and constructive criticisms. I am also thankful to other members who I have worked with during my time in the Johnson lab: Dr. Williard Werner, Dr. Erin Taylor, Dr. Myong-chul Koag, Jiawen Li, Brian Villalba, Tyler Dangerfield, and Ted Deng.

Finally, I would like to thank my dear wife, Min. She is always the source of love, concern, and encouragements. My words cannot fully express my appreciation for her understanding and support throughout my graduate study.

# Kinetic Studies of HIV-1 Reverse Transcriptase Nucleotide Selectivity, Drug Resistance and RNase H Activity

An Li, Ph.D.

The University of Texas at Austin, 2016

Supervisor: Kenneth Allen Johnson

Mechanisms of nucleotide selectivity and drug resistance by HIV-1 Reverse Transcriptase (HIVRT) were examined using rapid kinetic methods and global data fitting. Thymidine analog resistance mutations provide only two-fold discrimination against the incorporation of AZT-triphosphate; therefore, it is generally believed that resistance arises from nucleotide excision, where ATP reacts with the 3' terminal base to produce a dinucleotide tetraphosphate and a 3'-OH terminated primer capable of subsequent extension. Single turnover kinetic analysis with global data fitting revealed the intrinsic rate and equilibrium constants governing each step leading to resistance to chain terminators. Our data suggested the net resistance to AZT arises from nearly equal contributions involving discrimination during incorporation (2x), enhanced ATP-dependent excision (2-5x), reduced binding of the next correct nucleotide (2.6-5x) and more favored binding of the DNA primer in the N-site (5-20x).

Chemistry is generally the only rate-limiting step for product formation during DNA polymerization with a DNA template, but with an RNA template we show that pyrophosphate (PPi) release was rate-limiting. Due to the slow PPi release step, the rate

of reversal of chemistry could also be determined affording the first measurement of the equilibrium constants governing polymerization and the first complete free energy profile for HIVRT. Although PPi release is rate-limiting, nucleotide binding remains as the specificity-determining step. However, PPi release becomes exceedingly slow following mis-incorporation and thereby contributes to the specificity constant. Our data demonstrate that the fidelity of HIVRT has been underestimated by >20-fold in the past 20 years since the slow PPi release has been overlooked.

The rate-limiting PPi release allows synchronization and coordination of the polymerase and RNase-H activities. Studies were undertaken to examine proposed coordination between the polymerase and RNase-H activities. Direct, simultaneous measurement of the two activities established a mechanism by which polymerization and RNase-H activities are coordinated by working independently at comparable rates, but with different efficiencies. In contrast to polymerization, RNase-H requires ~4–6 base pairs extending beyond the active site for optimal reactivity, providing fast but infrequent cleavage events. Consequently, the polymerase and RNase-H activities are seamlessly coordinated without any direct communication between the two sites.

# Table of Contents

# List of Tables

# List of Figures

# List of Schemes

# Chapter 1: Introduction

## 1.1 PROJECT SUMMARY

In this thesis several projects were completed to understand the nucleotide selectivity and drug resistance by HIV-1 Reverse Transcriptase (HIVRT). Mechanistic perspectives of HIVRT were also explored in this dissertation. These studies provide insights into the design and evaluation of new drugs needed to manage HIV infections. Chapter 1 provides a comprehensive review of HIVRT. Chapter 2 reports that AZT drug resistance conferred by dipeptide insertions in the fingers of HIVRT involves multiple mechanisms that sum to define the net resistance. Chapter 3 explores effects of AZT resistance mutations and dipeptide insertions in the fingers of HIVRT on ATP-dependent excision of thymidine analogues. Chapter 4 reports that the polymerase and RNase H active sites of HIVRT work simultaneously and independently. Chapter 5 presents the rate-limiting pyrophosphate release by HIVRT and its effect on enzyme specificity. Chapter 6 explores effects of an R72A mutation on kinetics and fidelity of HIVRT.

## 1.2 AN OVERVIEW OF HIV

Since the discovery of Human Immunodeficiency Virus (HIV) as the causative agent of Aquired Immune Deficiency Syndrome (AIDS) in the early 1980s (*1*-*2*), HIV infection has been a severe public health problem over all the world. According to the 2014 report of the World Health Organization (WHO), 36.9 million people were living with HIV, with 2.0 million new infections occurring in 2014 (*3*). Without effective antiretroviral therapy, the immune system of an infected individual will be impaired gradually over time, making them more vulnerable to other infections, causing a series of

complications leading to death (*4*). Although the global HIV response has been remarkably transformed in the past 15 years and HIV treatment reached almost 16 million people in mid-2015 (*5*), there were still 1.2 million AIDS-related deaths in 2014 (*3*), highlighting the need for further research to find more effective antiretroviral therapies.

Knowledge of the molecular virology related to HIV infection is necessary for discussions of the development of antiretroviral treatment and accompanying challenges. HIV can be grouped into two major types, HIV-type 1 (HIV-1) and HIV-type 2 (HIV-2). It is clear that nearly all infections are due to the spread of HIV-1, whereas HIV-2 is restricted to West Africa and is less virulent (*6*). HIV is a lentivirus from the *Retroviridae* family, which has two identical copies of RNA genome enclosed in a capsid and a bi-layered lipid envelop (Figure 1.1).

Mature HIV particles have a spherical morphology of 100 nm in diameter, whose structure is shown in Figure 1.1 (*7*). The center of a HIV virion includes two copes of identical 9.7kb single stranded viral RNA (*8*) that encode structural proteins (*gag* and *env*), enzymes (*pol*), and other regulatory/accessory proteins (*tat*, *rev*, *vpr*, *vpu*, *vif*, and *nef*). The *gag* gene encodes the structural proteins that form the truncated cone-shaped capsid (p24) and the matrix (p17). The *pol* gene encodes enzymes necessary for viral replication, including reverse trancriptase that retrotranscribes the viral RNA genome into DNA, integrase that inserts linear double-stranded DNA into the host chromosome (provirus integration), and the protease that cleaves the Gag-Pol polyproteins into their components (*9*). The viral genomic RNA and enzymes are packed into the capsid, which

is surrounded by a bilayered membrane (or envelope) derived from the host cell's outer surface during the process of budding. The *env* gene encodes two major viral glycoproteins, gp41 and gp120 that are required for viral entry (*9*). The regulatory/accessory genes play key roles in modulating virus replication, whose functions are reviewed in (*10*).

The replication cycle of HIV (right panel, Figure 1.1) includes 1) virus entry, involving CD4 binding, co-receptor (CXCR4 or CCR5) interactions and membrane fusion (step 1-3); 2) post-entry events, including reverse transcription that converts viral RNA into proviral DNA (detailed in section 1.3), nuclear import, and integration of proviral DNA into host genome (step 4-5); 3) gene expression, where the integrated viral genome is expressed using the host machinery (step 6); 4) virus particle production, in which a Gag precursor polyprotein assists plasma membrane targeting and binding (step 7) (*11*), then RNA is encapsidated and final assembly and budding occur (step 8). 5) maturation, where protease cuts the precursor polyproteins to generate mature Gag and Pol proteins (step 9) (*12*) (reviewed in (*7, 9, 13*)).

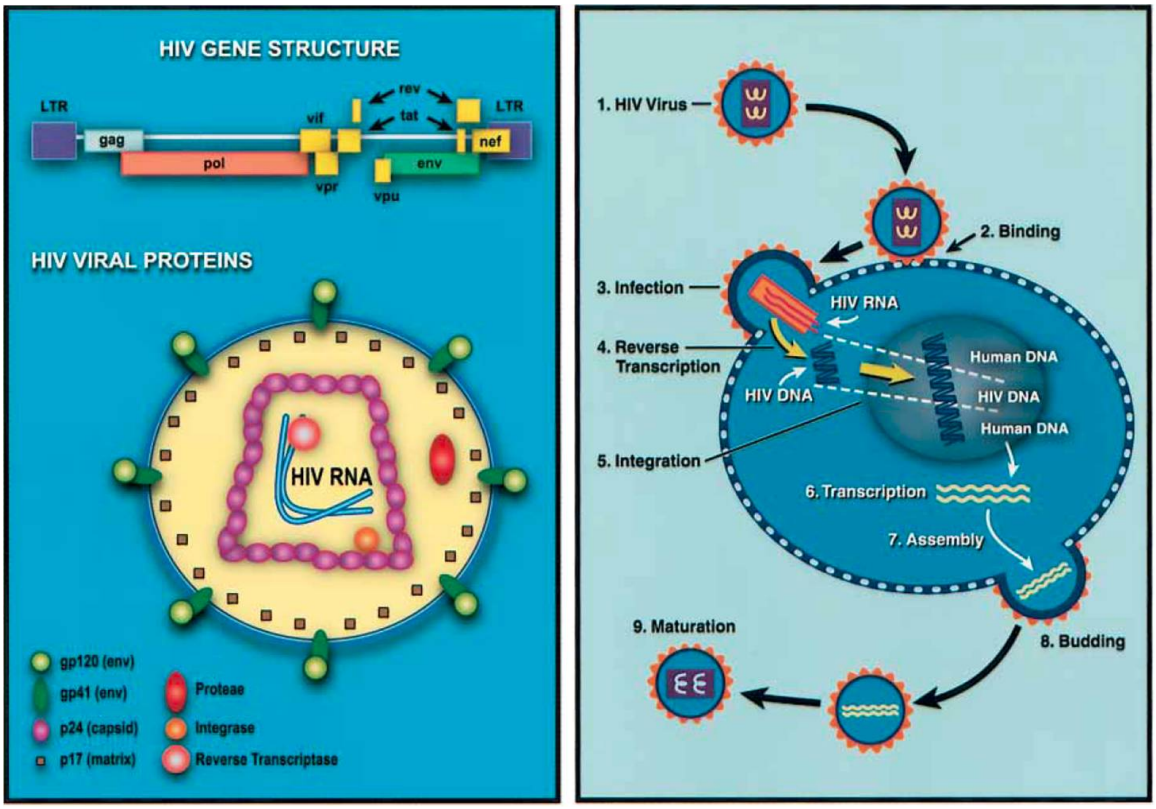



**Figure 1.1** *(Left) HIV viral proteins and gene structure. (Right) Replication of HIV*. Reproduced from Chinen and Shearer *(7)*. (Left) HIV gene structure is shown on the top. LTR: long terminal repeat. The *gag-pol* gene encodes two precursor polyproteins that are cleaved to generate structural proteins and 3 viral enzymes: reverse transcriptase (RT), integrase (IN), and protease (PR). The *env* gene encodes the gp160 polypeptide precursor that contains two glycoproteins, gp 120 and gp41. HIV accessory genes are shown in yellow. (Right) HIV infection begins with binding (step 2), then viral core is released into the cell (step 3). Viral RNA is then converted to viral cDNA (step 4) that can be inserted into host genome (step 5). The viral DNA is transcribed into mRNA (step 6), followed by protein synthesis and assembly (step 7). Finally, new mature viruses are released from the host cell (step 8-9).

## 1.3 HIV RT FUNCTION IN VIRAL REPLICATION

The multifunctional reverse transcriptase initiates the process of reverse transcription once the viral RNA genomes are released into the cytoplasm of infected cells. Although some viral and cellular factors may assist in the process of reverse transcription, HIV-1 RT contains all necessary enzymatic functions for the conversion, including RNA- and DNA-dependent DNA polymerase and RNase H activities.

The individual steps for provirus DNA synthesis are illustrated in Figure 1.2. There is an 18-nucleotide-long primer binding site (PBS), located approximately 180 nucleotides from the 5' end of the viral genome (*14*). The host $tRNA^{lys3}$ is base paired to the PBS via sequence complementarity and serves as the primer for (-) strand DNA synthesis using viral RNA as the template until it stops at the 5' end of viral genome. The resulting DNA/RNA hybrid is a substrate for RNase H that selectively removes the copied RNA strand, exposing the newly synthesized DNA strand. Since the viral genome has two copies of identical sequence at the 5' and 3' ends (R or repeat), the nascent DNA is able to anneal with the 3' end of one of the two genomic RNA strands so that DNA synthesis can resume, a process called '(-) strand DNA transfer'. As DNA synthesis proceeds, RNase H continues to hydrolyze the RNA strand with two exceptions known as polypurine tracts (cPPT and 3' PPT) that are proposed to adopt a special conformation (*15*) and therefore are resistant to RNase H cleavage. The 3' PPT serves as the primer for the initiation of the (+) strand DNA synthesis that will continue until it encounters a modified A in the template, which is the $19^{th}$ base from the 3' end of the tRNA primer. After that, PPT and tRNA primers are removed by RNase H cleavages and the second (or

plus) strand transfer occurs based on homology between (-) and (+) strand DNA primer binding sites. Then, bidirectional DNA synthesis extends both the (-) and (+) strands to the ends of both templates (reviewed in (*14, 16-17*)).

The role of RNase H in reverse transcription has been reviewed by Sharon and James (*18*). The polymerization-dependent RNase H is presumably responsible for generating nicks on genomic RNA but is not sufficient to completely remove the copied RNA template since the polymerization rate is much greater than the cleavage rate of the enzyme (*19-22*). Consequently, the synthesis-independent RNase H cleavages take place to remove genomic RNA that remains annealed to the (-) strand DNA (*19, 21*). Polymerization-independent hydrolysis also participates in the generation of PPT primer and the removal of extended tRNA and PPT primers (*23-24*).

It has been suggested that HIV-1 RT is an error-prone DNA polymerase, which accounts for the rapid emergence of mutations to evade the immune system and antiviral drugs. Although the fidelity of HIV-1 RT measured *in vitro* is particularly low (*25-26*), the mutation rates of HIV-1 replication measured *in vivo* is similar to other retroviruses (*27-29*), and the observed diversity of viral sequences is due to the rapid replication of HIV-1 viruses in a large number of infected cells (*14*). As another component of the viral replication cycle, the relative contribution of host DNA-dependent RNA polymerase II (RNA pol II) to the overall viral mutation rate is still unknown. Nonetheless, the ability of HIV-1 to diverge rapidly via reverse transcription is important for the virus to become resistant to all approved antiretroviral therapies.

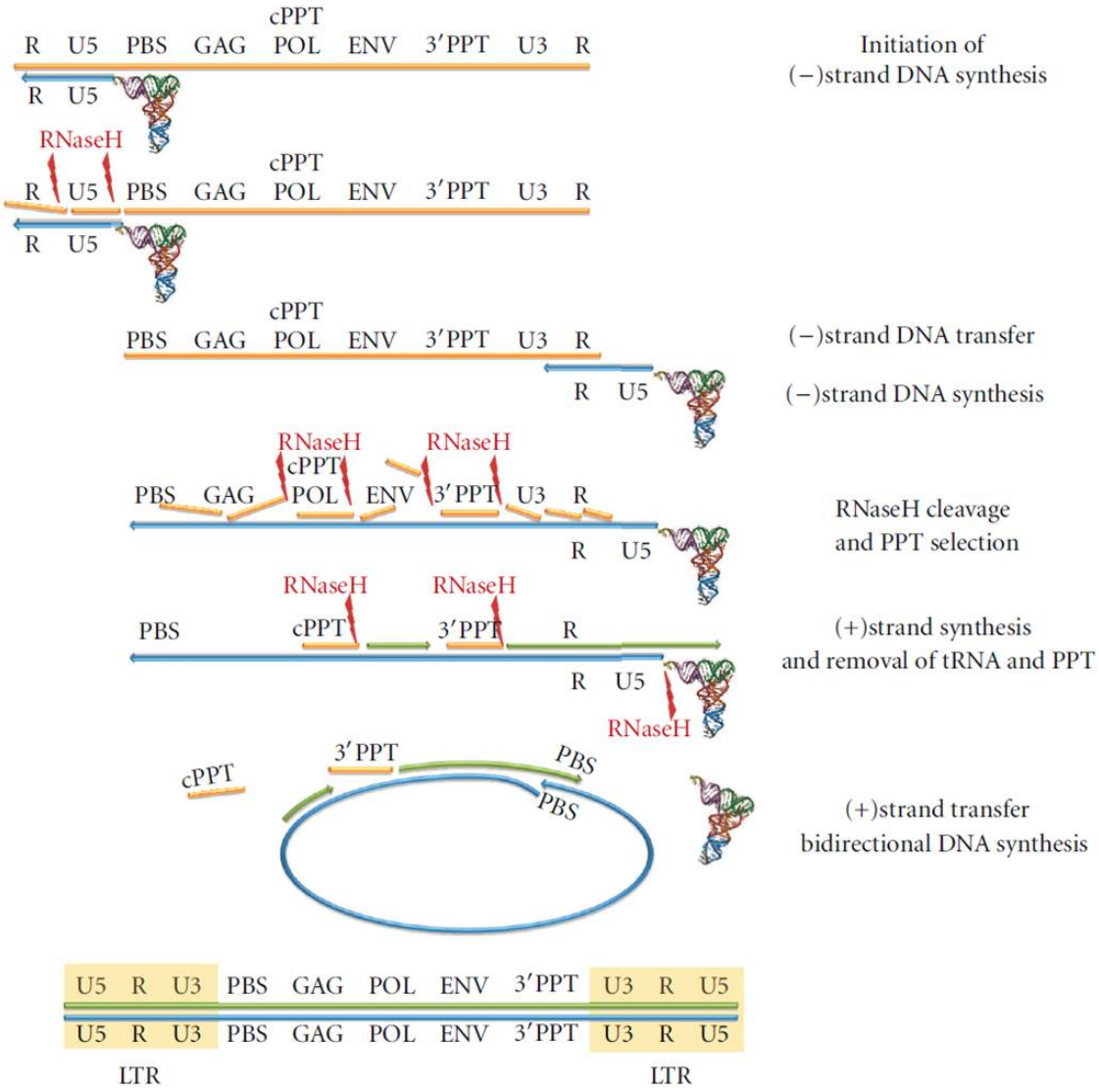



**Figure 1.2** *HIV-1 Reverse Transcription Process*. Reproduced from Esposito, Corona and Tramontano(*16*). Step 1: Host $tRNA^{lys3}$ is hybridized to the primer binding site (PBS) and initiates (-) strand DNA synthesis. Step 2: RNase H degrades the copied RNA segment in a DNA/RNA hybrid and frees the newly synthesized DNA strand. Step 3: The first strand transfer occurs since the viral genome has identical sequence (R or repeat) at the two ends, after which (-) strand DNA synthesis continues. Step 4: As DNA synthesis proceeds, RNase H degrades the RNA strand, leaving two polypurine tracts (PPTs) intact, namely, the central PPT (cPPT) and 3'PPT. Step 5: The latter serves as the primer for (+) strand DNA synthesis, which stops at a modified A of the 3'-end of the tRNA. Then, tRNA is removed by RNase H. Step 6: The second strand transfer occurs based on homology between (-) and (+) strand DNA primer binding sites. Then, bidirectional DNA synthesis completes DNA synthesis.

## 1.4 THE STRUCTURE OF HIV RT

HIV-1 RT is a heterodimer, consisting of two related chains, p66 and p51. The p66 subunit is composed of the polymerase and RNase H domains and fulfills the catalytic role, while the p51 subunit is derived from p66 by proteolytic cleavage to remove the C-terminal RNase-H domain (*30*). The polymerase domain shows a conserved right hand morphology and consists of the fingers (residues 1-85 and 118-155), palm (residues 86-117 and 156-237), thumb (residues 238-318), and connection subdomains (residues 319-426) (*31-32*). The RNase H domain comprises residues 427-560 (Figure 1.3). On the other hand, p51 adopts a different conformation after a complex maturation process (*33-34*) and is proposed to only provide structural support (*31-32*) and form part of the nucleic acid binding cleft.

### Active Sites

The polymerase active site was described in detail by Huang *et al.* (*35*) and key residues involved in nucleotide binding, metal ions coordination, and catalysis have been reviewed (*36*). Notably, three aspartatic acid residues (D110, D185 and D186) form the catalytic triad and help to position the two divalent ion cofactors ($Mg^{2+}$). R72 and K65 of the fingers interact with the β- and γ- phosphates of the incoming dNTP, respectively (Figure 1.3, bottom left). The importance of R72 in coordinating the phosphate was supported by the fact that a R72A mutation significantly impaired the PPi release step (*37*). Residue Y115 serves as the 'steric gate' that contributes to discriminate against the ribonucleoside triphosphates. A two metal-ion mechanism is utilized for catalysis (*38*), in which metal ion A coordinates the 3'-hydroxyl group of the primer and lowers its pKa,

whereas metal ion B neutralizes the negative charge of the incoming nucleotide and stabilizes the transition state. Interestingly, recent structural data obtained using time-resolved X-ray crystallography have suggested the transient existence of a third $Mg^{2+}$ on the opposite side of the A- and B-site $Mg^{2+}$ ions. This extra $Mg^{2+}$ may appear at the active site during the process of phosphodiester bond formation, stabilizing the transition state (*39-40*) or arrive after catalysis and promote the pyrophosphorolysis activity (*41*). Notably, the third metal binding site is not formed during incorrect nucleotide insertion, which might account for the inability of the polymerase to perform the reverse reaction when a mismatched primer terminus was present as shown Chapter 5.

The RNase H active site is ~60 Å away from the polymerase active site and comprises four conserved carboxylates residues (D443, E478, D498, and D549) (*18*) that can coordinate two divalent $Mg^{2+}$ (Figure 1.4 bottom right). Although both one- (*42*) and two- (*38*) metal facilitated RNA hydrolysis have been suggested, more recent research data have supported the two metal-ion mechanism (*43-44*). In this mechanism, the first metal ion A may assist nucleophilic attack by bringing the nucleophile close to the scissile phosphate, whereas the second metal ion B stabilizes the transition state and facilitates leaving of the 3' oxyanion group (*43, 45*).

**Nucleic Acid Binding Cleft**

In 1992, a structure of HIV-1 RT/DNA complex at 7 Å resolution showed that the nucleic acid binds in a cleft on the surface of the enzyme (*46*). The cleft was then observed in the crystal structure at 3.5 Å resolution of HIV-1 RT complexed with an inhibitor, into which a duplex A-form DNA/RNA hybrid can be modeled (*31*). The

nucleic acid binding cleft of HIV-1 RT is formed by the polymerase and RNase H domains of the p66 subunit, whereas the floor of the cleft is formed by the thumb and connection subdomains of the p51 subunit (*36*). The cleft runs between the RNase H and polymerase active sites, which is able to accommodate 17 to 18 base pairs, depending on the nucleic acid substrate (*15, 32*). Most RT-nucleic acid interactions involve the sugar-phosphate backbone and therefore are non-specific. There are extensive interactions between conserved structural motifs of RT and the primer/template near the polymerase active site that play important roles in holding and precisely positioning the nucleic acid relative to the polymerase site. These motifs include the 'template grip' comprising V75, R78, N81, E89, P157, and G93 of the p66 fingers (*15, 47*), as well as the 'DNA primer grip' consisting of the β12-β13 hairpin in HIV-1 RT (*48*). Similarly, the 'RNase H primer grip' in the p66 connection domain plays a critical role in positioning an RNA-DNA hybrid for specific cleavage (*49-50*). Additional interactions of the RNA template with the 'template grip' (via 2'-OH) were observed (*51*), which may account for the higher fidelity of RNA-dependent DNA polymerization compared to DNA-dependent DNA polymerization by RT (*52-53*). However, no structural element of RT was found that could systematically guide the RNA strand to the RNase H active site (*51*).

The geometry of a DNA duplex bound to HIV RT starts with the A-form at the polymerase active site, but soon switches to B-form after a ~40° bend adjacent to the p66 thumb (*32*). Besides this typical kink near p66 thumb, the DNA/RNA hybrid bound to RT showed an additional 20-30° bend when entering the RNase H domain (*54*), and the hybrid structures are mostly in the A form (*51, 54*), with one exception where the (PPT)

RNA/DNA adopts an intermediate between A- and B-form, which was termed the 'H-form' (*15*). To date, there are only three structures of HIV-1 RT that are in complex with a DNA/RNA hybrid, but none of the bound hybrid is captured in the RNase H competent mode; therefore, the mechanism by which an RNA template reaches the active site for degradation is still unknown (*15, 51, 54*).

**Conformational Dynamics during DNA Synthesis**

HIV-1 RT undergoes extensive conformational rearrangements during different stages of DNA synthesis and crystal structures corresponding to some of the states are now available, which has guided the design of biochemical assays to explore these structural transitions (*36*). A large conformational change of the p66 thumb was identified by comparing structures of unliganded HIV-1 RT (*55-56*) and RT/DNA duplex (*32*). The thumb subdomain is in a closed configuration relative to the fingers in unliganded RT structures and therefore blocks the nucleic acid binding cleft. However, binding of nucleic acid substrate moves the thumb subdomain away from the fingers to accommodate the primer/template. At 1998, Huang *et al.* (*35*) successfully trapped the catalytic complex of HIV-1 RT by chemical cross-linking of the DNA and using a dideoxy-terminated primer with the next correct base in order to obtain the structure of RT-primer/template-dNTP ternary complex in a closed form. This is the first time an RT structure was obtained with DNA bound. Structures in the presence and absence of the dNTP then revealed nucleotide-induced conformational changes from an *open* to a *closed* state. An overlay of RT structures in the *open* (1J5O (*57*)) and *closed* (1RTD (*35*)) states are shown in Figure 1.4 to illustrate the global changes in structure occurring in the

enzyme-DNA duplex upon dNTP binding. The fingers domain undergoes the largest conformational change (especially the β3-β4 loop), where several positively charged amino acids close down on the triphosphate portion of the incoming dNTP to facilitate catalysis (*58*). Results from the MD simulation using Elber's Milestoning method provided molecular details of the conformational transition and helped to explain how conformational dynamics of DNA polymerases select correct substrates (*58*). However, there is no direct structural information in HIV-1 RT to illuminate the chemistry and PPi release steps during nucleotide incorporation. After PPi dissociation, RT must translocate on the nucleic acid substrate to free the occupied nucleotide-binding site (N site), the molecular details of which are still unknown. But structures of HIV-1 RT trapped at pre- and post-translocation states suggest that the conserved YMDD loop may be involved in translocation by 'acting as a springboard to propel the primer terminus from the N to the P site after dNMP incorporation' (*59*), but there is not data to support such a model.

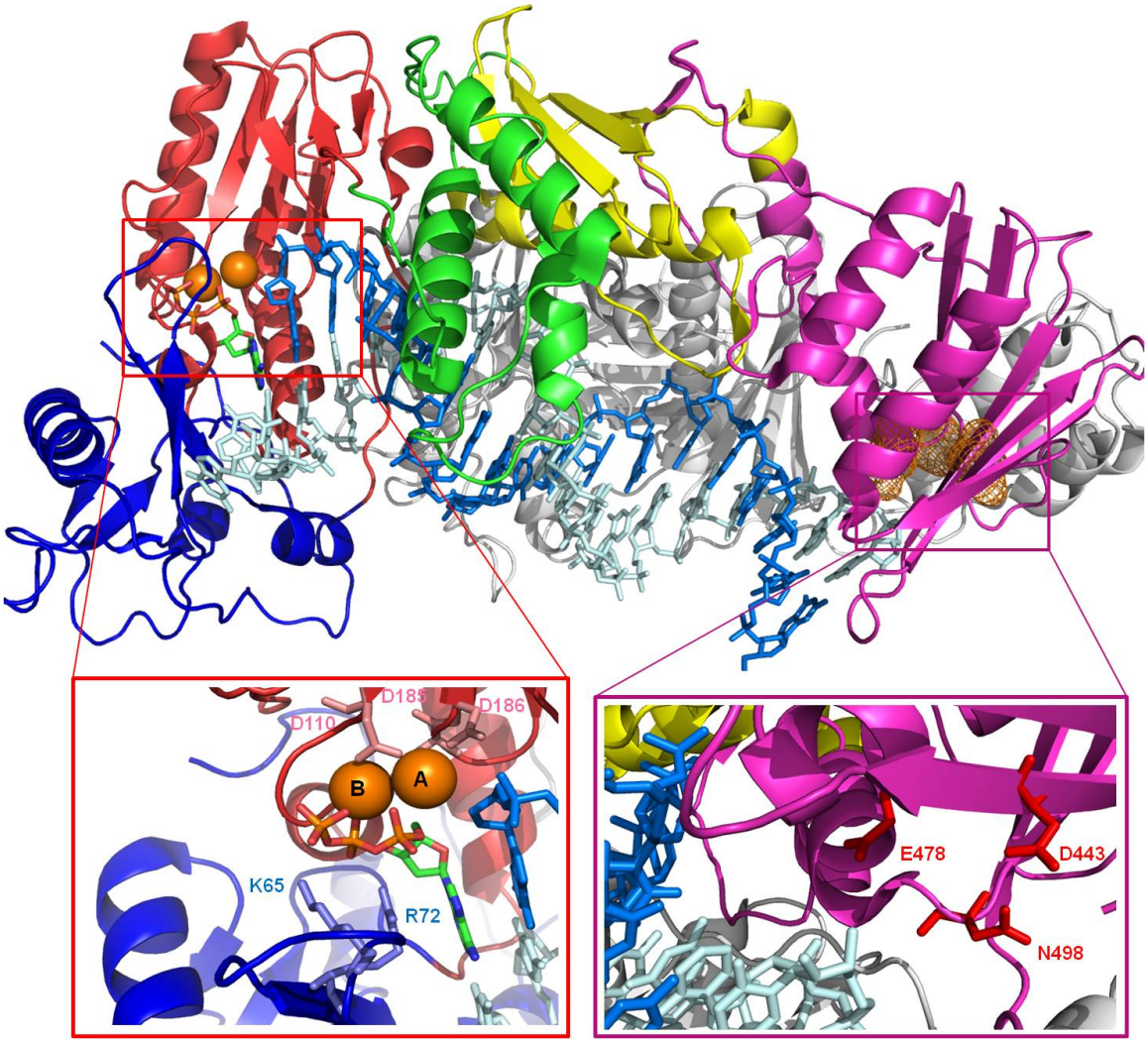



**Figure 1.3** *The Crystal Structure of HIV-1 RT in complex with a DNA/RNA Hybrid (PDB: 4PQU).* (Upper panel) The enzyme has two subunits: the p66 (colored) and p51 (gray). Structure of the polymerase domain resembles a right hand, consisting of fingers (blue), palm (red), thumb (green), and connection (yellow) subdomains. The RNase H domain (magenta) is ~60 Å away from the polymerase active site. $Mg^{2+}$ ions are shown as orange spheres. The incoming nucleotide is dATP (green). The DNA primer and RNA template are shown in blue and cyan, respectively. The polymerase and RNase H active sites are highlighted by red and magenta boxes and zoomed in the bottom. (Bottom left) The polymerase active site is located in the middle of palm, fingers, and thumb subdomains. The catalytic triad (D110, D185 and D186) are shown in pink. Basic residues (K65 and R72) interacting with the phosphates of dATP are shown in blue. (Bottom right) The RNase H active site contains the carboxylates residues D443, E478, D498 (D498N mutation was made to obtain this structure), and D549 (missing in this structure).

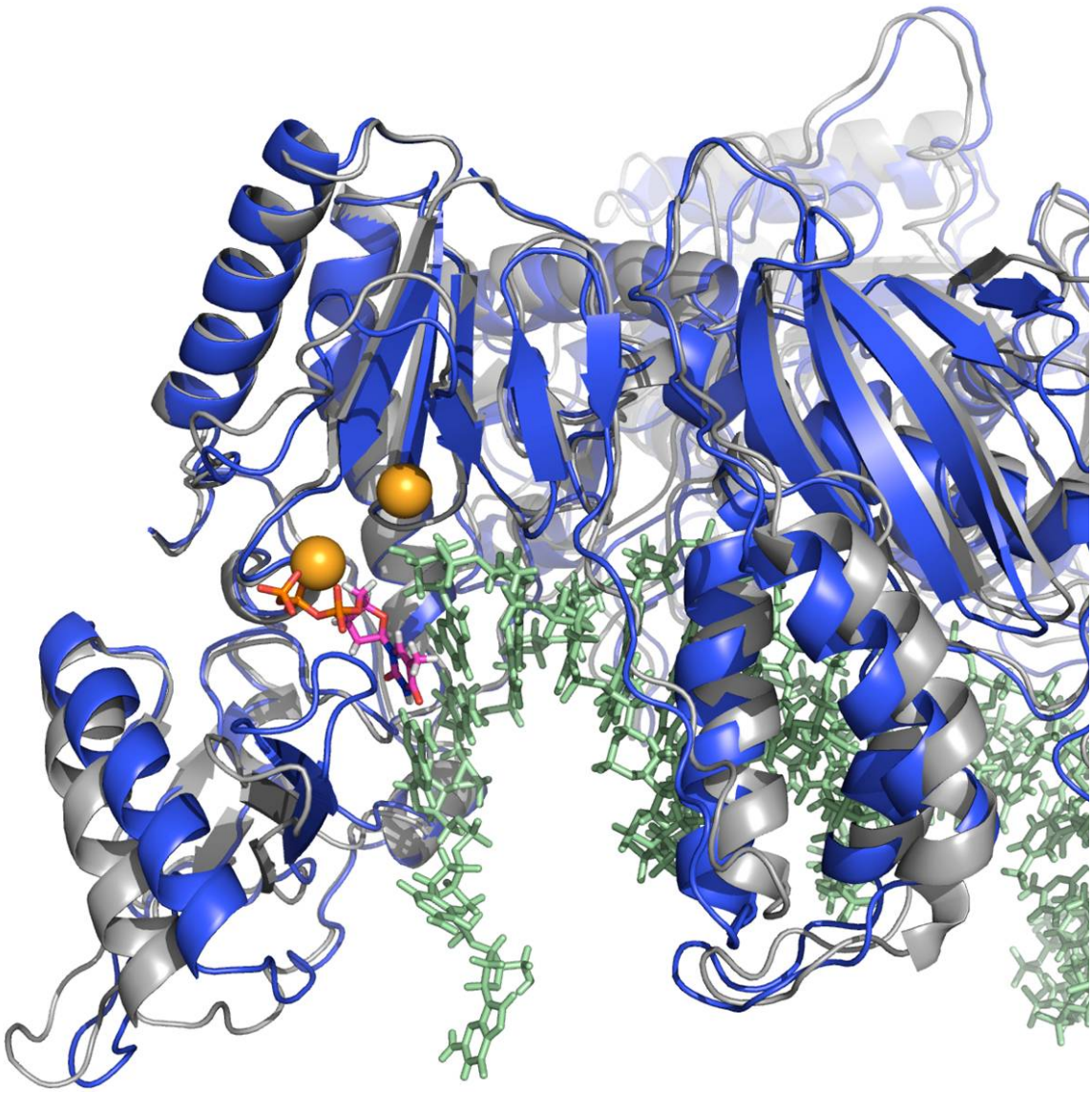


**Figure 1.4** *Nucleotide-induced Conformaiton Changes* (*58*). Overlay of HIVRT structures 1RTD (*35*) and 1J5O (*57*) is present to illustrate the global changes in structure occurring in the enzyme-DNA duplex from the open (gray) to closed (blue) states (emphasis is put on the polymerase active site). The P51 subunit is not shown for clarity. Incoming dNTP substrate (purple-orange) and two $Mg^{2+}$ ions (yellow) are shown together with the primer (palecyan)/template (forest green) duplex. The fingers domain undergoes the largest conformational change, where several positively charged amino acids come into contact with the triphosphate portion of the incoming dNTP to facilitate catalysis.

## 1.5 HIV RT AS A DRUG TARGET

Six categories of antiretroviral therapies (ARTs) exist, targeting different stages during HIV infection lifecycle: nucleoside/nucleotide analogue RT inhibitors (NRTIs), non-nucleoside RT inhibitors (NNRTIs), protease inhibitors (PIs), integrase inhibitors (IIs), fusion inhibitors (FIs), and viral entry inhibitors (EIs). As mentioned before, HIV-1 RT plays a central role in viral replication and therefore has been a major target for therapeutic treatment of HIV infections (*60-61*). Among the 26 drugs approved to treat HIV-1 infections, 14 are RT inhibitors (*14*), consisting of two groups of compounds (NRTIs and NNRTIs) that function with distinct mechanisms.

### NRTIs and the Drug Resistance

NRTIs are structurally diverse analogues of natural nucleotides that lack the 3'-OH group and therefore can serve as chain terminators upon incorporation (Figure 1.5). These agents are inactive and must be phosphorylated by cellular kinases to their active form, the triphosphate derivatives (*62*). They inhibit the reverse transcription by the same mechanism: the activated triphosphate form of the compounds incorporate into viral DNA and prevent the subsequent DNA synthesis due to the lack of 3'-OH group. It is important that the eukaryotic DNA polymerases are able to discriminate against these compounds to minimize toxic side effects. However, it was found that kinetics of incorporation of nucleoside analogues by the human mitochondrial DNA polymerase (Pol γ) is correlated with toxic side effects of NRTIs (*63*). Theoretically, NRTIs should not be toxic since human mitochondrial DNA polymerase has a proof-reading activity and exhibits moderately high fidelity during DNA replication, whereas RT is an lower fidelity

polymerase with less discrimination against these analogues as substrates. However, not all NRTIs can be efficiently removed by the proofreading exonuclease of the human mitochondrial DNA polymerase (*63*). A parameter termed the therapeutic index (the ratio of the discrimination by RT divided by the discrimination by Pol γ) was introduced to assess the effectiveness of NRTIs against their toxicity (*64*). Notably, two nucleoside analogues (ddC and d4T)) exhibiting lowest therapeutic index are no longer used in clinics because of their high toxicity (reviewed in (*64*)).

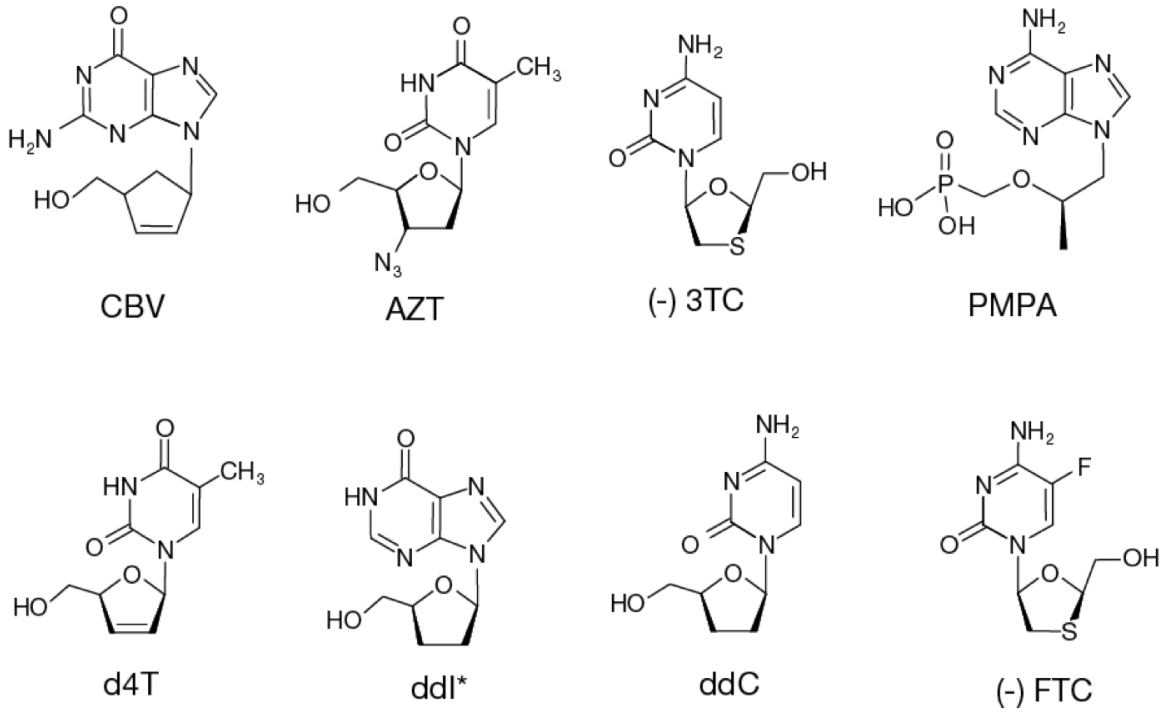



**Figure 1.5** *Structures of NRTIs Approved by the FDA.* CBV-abacvir, AZT-zidovudine, 3TC-lamivudine, PMPA (TFV)-tenofovir, d4T-stavudine, ddI-didanosine, ddC-zalcitabine, and FTC-emtricitabine. PMPA is a nucleotide analogue. TFV and CBV are given as pro-drugs (not shown), whereas ddI is metabolized to ddA (*65*).

Appearance of drug resistance represents another major limitation for the use of NRTIs to treat HIV-1 infections. With the availability of highly active antiretroviral therapy (HAART), HIV infection is now controllable as a chronic disease in patients who have access to medication and many of whom achieve a sustained viral suppression, but this requires long-term medication (*66*). Prolonged use of NRTIs inevitably results in the appearance of viruses resistant to these compounds (*67*). Some mutations only confer resistance to one of the NRTIs, whereas certain combinations of mutations (e.g., the 69 insertion complex, Q151M complex and thymidine analogue mutations (TAMs)) make RT resistant to all or almost all NRTIs approved by the US FDA (Figure 1.6) (*68*). HIV-1 RT mutants become resistant to NRTIs by two mechanisms: i) enhanced discrimination against NRTIs, which means that RT restricts incorporation of the analogues and therefore reduces chain termination; and ii) enhanced excision of chain terminators from the 3'-end of the blocked primer, so new DNA synthesis can resume (*69*). In nucleotide-dependent excision, ATP reacts with the 3' terminal base to produce a dinucleotide tetraphosphate (Ap4N, where N is the excised base) and a 3'-OH terminated primer capable of subsequent extension. This reaction is analogous to the normal reversal of chemistry in which pyrophosphate reacts to re-synthesize dNTP, but in this case, ATP provides the reactive pyrophosphate portion. In this thesis, both mechanisms of resistance are investigated to more accurately quantify their contributions to net resistance.

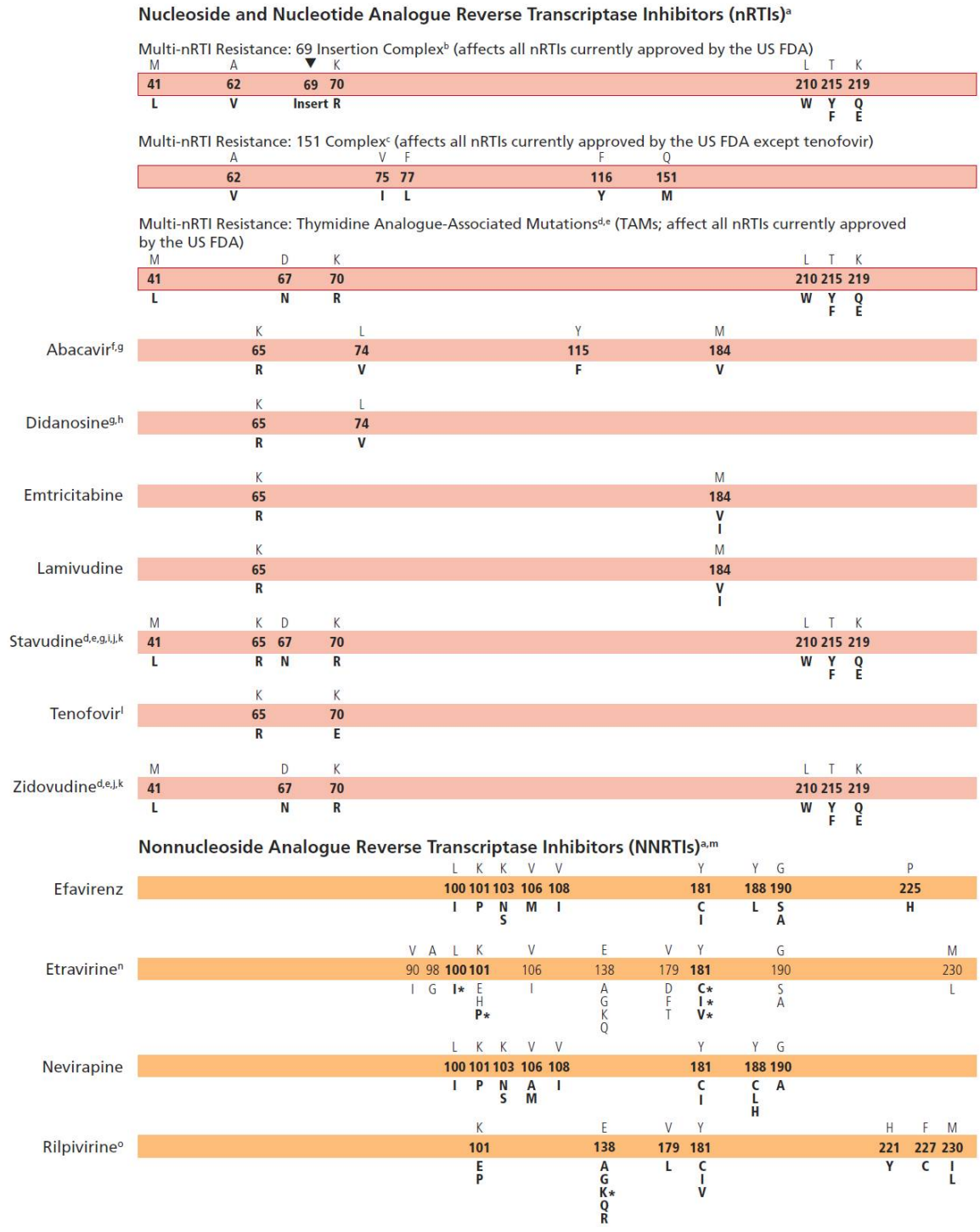

Nucleoside and Nucleotide Analogue Reverse Transcriptase Inhibitors (nRTIs)[a]
Multi-nRTI Resistance: 69 Insertion Complex[b] (affects all nRTIs currently approved by the US FDA)
M A ▼ K L T K
41 62 69 70 210 215 219
L V Insert R W Y Q
F E
Multi-nRTI Resistance: 151 Complex[c] (affects all nRTIs currently approved by the US FDA except tenofovir)
A V F F Q
62 75 77 116 151
V I L Y M
Multi-nRTI Resistance: Thymidine Analogue-Associated Mutations[d,e] (TAMs; affect all nRTIs currently approved by the US FDA)
M D K L T K
41 67 70 210 215 219
L N R W Y Q
F E
Abacavir[f,g]
K L Y M
65 74 115 184
R V F V
Didanosine[g,h]
K L
65 74
R V
Emtricitabine
K M
65 184
R V
I
Lamivudine
K M
65 184
R V
I
Stavudine[d,e,g,i,j,k]
M K D K L T K
41 65 67 70 210 215 219
L R N R W Y Q
F E
Tenofovir[l]
K K
65 70
R E
Zidovudine[d,e,j,k]
M D K L T K
41 67 70 210 215 219
L N R W Y Q
F E
Nonnucleoside Analogue Reverse Transcriptase Inhibitors (NNRTIs)[a,m]
Efavirenz
L K K V V Y Y G P
100 101 103 106 108 181 188 190 225
I P N M I C L S H
S A I
Etravirine[n]
V A L K V E V Y G M
90 98 100 101 106 138 179 181 190 230
I G I* E I A D Y* S L
H F I* A
P* K C* A
T V*
Nevirapine
L K K V V Y Y G
100 101 103 106 108 181 188 190
I P N A I C C A
S M L H
Rilpivirine[o]
K E V Y H F M
101 138 179 181 221 227 230
E A L C Y C I
P G I L
K* V
Q
R

**Figure 1.6**

**Figure 1.6** *Mutations in the Reverse Transcriptase Gene Associated with Resistance to RT Inhibitors.* Reproduced from Johnson *et al. (68)*. (Top) Mutations associated with the resistance to NRTIs are shown on the top (pale pink bars), including the 69 insertion complex, 151 complex, TAMs, and the mutations associated with drug resistance to a specific inhibitor. (Bottom) Mutations associated with the resistance to NNRTIs are shown on the botthom (orange bars). Single-letter amino acid codes at the top of colored bars represent wild-type residues, whereas those beneath the bars are amino acid substitutions conferring resistance. Numbers in the bars are the positions of substituted residues. Black triangle denotes the insertion sites. Superscript letters near the compound names represent user notes, and please refer to the text of the reference for details.

K65R and M184V are two representative mutations for the discrimination mechanism. In general this class of mutations are located in close proximity to the incoming nucleotide and therefore interfere with its binding. K65R has an enhanced interaction with the side chain of R72 and forms a molecular platform that restricts the adaptability of the polymerase active site and causes decreased rates for substrate incorporation and NRTIs excision (*70*). This might be the structural basis for the mutual phenotypic antagonisms between the K65R and TAMs mutants, which likely accounts for the fact that these two kinds of mutations rarely coexist (*71*). K65R/A substitutions may also involve in the enhanced polymerase fidelity of RT by decreasing the dNTP misinsertion and mispaired primer extension efficiencies (*72*). A single mutation of residue 184 (M184V/I) is sufficient to confer high levels of resistance to 3TC. Comparisons of the structures of HIV-1 RT duplex (with bound DNA) (*32*) and ternary complexes (with bound dNTP) (*35*) with M184V and M184I mutants (*57, 73*), together with the modeling of the inhibitor into the active site of these structures have suggested a steric clash between the sugar moiety of 3TC-TP and the mutation at position 184, which might account for the diminished incorporation of 3TC. While in the case of tenofovir (TFV) resistance, M184V seems to offset the extent of resistance conferred by K65R (*74*). In addition, it was reported that M184V also played an important role in the excision mechanism and was able to severely compromise the removal of chain-terminating nucleotides, such as AZTTP (*75*).

"Thymidine analogue mutations" (TAMs) comprising M41L, D67N, K70R, L210W, T215Y/F and K219E/Q are a well studied example for the excision mechanism.

Two questions are important to understand the underlying mechanism for the resistance conferred by this complex of mutations: i) how do these mutations enhance the excision, and ii) why is the resistance relatively specific for AZT (*17*)? With regard to the first question, all of the mutations have no direct interactions with the AZTTP moiety of Ap4AZT (the dinucleotide tetraphosphate product of ATP mediated excision) (*76*), suggesting that AZT resistant mutations may not contribute to the discrimination against AZTTP. Indeed, a recent pre-steady-state kinetics study revealed that the TAMs mutations led to reduced rate of AZT incorporation, but that was offset by a lower $K_m$, yielding only moderate level of discrimination (*77*). Alternatively, it has been proposed that the mutations create a new binding site for ATP (*76*), where the pyrophosphate moiety of ATP is properly positioned, presumably leading to a tighter ATP binding that may facilitate the excision. However, there is no evidence to support the enhanced ATP-binding affinity (*78*), implying that an alternative mechanism may account for the enhanced excision efficiency. As for the second question, steric constraints involving the azido group cause the 3' end of an AZTMP terminated primer to preferentially reside at the nucleotide binding site (N site), favoring the excision (*79*). In fact, unpublished single molecule data from this lab have suggested that the DNA primer terminated with AZTMP shifts the DNA translocation to favor the pre-translocation complex (at the N site) and inhibits the nucleotide-induced isomerization. These two features favored the removal of AZTMP by ATP-mediated pyrophosphorolysis. However, these data failed to fully address this question since other analogues with azido group exhibited much lower excision efficiencies (*80*). So this question is going to be explored in further work.

**NNRTIs**

Nonnucleoside RT inhibitors (NNRTI) are a group of structurally diverse, small hydrophobic agents specifically targeting HIV-1 RT. Transient kinetic studies revealed that NNRTIs affected the nucleotide incorporation by allosterically inhibiting the rate of the chemical reaction (*81*). Structural studies of HIV-1 RT have shown that the NNRTI-binding pocket (NNIBP) only exists on the p66 subunit of the heterodimer upon NNRTI binding, which causes a series of short-range and long range distortions of the RT structure, including the precise geometry of the polymerase active site, especially the 'YMDD' motif (*82*), and the structural elements comprising the 'primer grip' which are involved in the proper positioning of the DNA primer (*55*). In addition, there are data suggesting that the binding of NNRTIs may constrain the mobility of the thumb subdomain thus slowing or preventing the primer/template translocation, which in turn inhibiting further extension of the nascent DNA strand (*31*). The combination of all these effects may account for the relatively high effectiveness of NNRTIs. However, the emergence of NNRTI resistance associated mutations (most of which are found in or around the NNIBP) is also the major constraint for their clinic use (Figure 1.6). Recent research data indicate that the drug resistance associated mutations are stimulated by NNRTIs at the early stages of replication (*83*).

The advent of multi-drug combination antiretroviral therapies has saved a lot of lives in developed countries. However, the toxicity and development of drug resistance are the major limitations of the highly active antiretroviral therapy (HAART). In the future, molecular modeling, biochemical, and enzymatic approaches should be combined

to explore the structure-function relationship of target enzymes, as well as the mechanisms for drug resistance, which would shed light on the development of new antiretroviral drugs.

## 1.6 KINETIC CHARACTERIZATION OF HIV RT

### Nucleotide-induced Conformational Change and Enzyme Specificity

A simplified model for polymerization was proposed more than 20 years ago (Scheme 1.1) for quantification of polymerase specificity using single-nucleotide incorporation kinetic measurements (reviewed in (*84*)). In this model, the nucleotide concentration dependence of incorporation rates defines the apparent dissociation constant of dNTP ($K_d$) and the maximum incorporation rate ($k_{pol}$). This model is valid when nucleotide binding is a rapid equilibrium process and chemistry is the only rate-limiting step for product formation. The two assumptions are largely valid since PPi release and translocation are generally fast. Consequently, the measurements of $k_{pol}$ and apparent $K_d$ provided an estimate for the specificity constant, $k_{cat}/K_m = k_{pol}/K_d$. However, nucleotide binding is not a simple, one-step process, as structural analyses reveal a large conformational change of the fingers domain that forms part of the nucleotide-binding pocket (*32, 35*). As shown in Scheme 1.2, binding of nucleotide involves a weak ground state binding ($1/K_1$), followed by an isomerization of the fingers domain from open ($ED_ndNTP$) to closed ($FD_ndNTP$). The role for this change in enzyme specificity, however, has been controversial (reviewed in (*85*)), as least partially, to the lack of a proper signal to directly measure the kinetics of the isomerization step. By site-specifically labeling HIV-1 RT with an environmentally-sensitive fluorophore (7-

diethylamino-3-((((2-maleimidyl) ethyl) amino) carbonyl) coumarin, or MDCC) on the fingers of the polymerase, Kellinger and Johnson (*77, 86*) were able to measure both the rate and the equilibrium constants for the nucleotide-induced conformation change, leading to the conclusion that the isomerization step governs specificity and analogue discrimination by HIV-1 RT. A similar conclusion was also obtained to explain substrate specificity of the T7 DNA polymerase (*87*). Although the simplified model is still valid in most situations and allows adequate quantification of specificity, the complete model is far more robust in terms that it is capable of revealing the mechanistic basis for specificity. In this dissertation, a comprehensive set of experiments and global data analysis were performed to explore the mechanistic basis for the observed discrimination against AZT by the 69 insertion complex (Chapter 2).

**Scheme 1.1** *The Simplified Model of Nucleotide Incorporation*

$$ED_n + dNTP \underset{}{\overset{K_d}{\rightleftharpoons}} ED_ndNTP \overset{k_{pol}}{\rightleftharpoons} ED_{n+1}.PPi \xrightarrow{fast} ED_{n+1} + PPi$$

**Scheme 1.2** *The Complete Model of Nucleotide Incorporation*

$$ED_n + dNTP \overset{K_1}{\rightleftharpoons} ED_ndNTP \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} FD_ndNTP \overset{k_3}{\rightleftharpoons} FD_{n+1}.PPi \xrightarrow{fast} ED_{n+1} + PPi$$

**Role of Pyrophosphate Release**

PPi release is generally assumed to be fast based on the fact that during processive DNA synthesis, there is no significant delay after the incorporation of one nucleotide and

preceding the binding and incorporation of the next (*20, 88*). The fast PPi release is critical for understanding specificity, but kinetics of PPi dissociation has never been measured for HIV-1 RT. There have been a few instances in previous studies of the human mitochondrial DNA polymerase (*89-91*) where PPi release was much slower than the chemistry step during the incorporation of modified or mismatched nucleotides, leading to altered nucleotide specificity constants. In addition, the slow PPi release was proposed to provide a novel type of proofreading mechanism by allowing the reversal of chemistry (*89-90*). Previous measurements of the rates of polymerization and pyrophosphate release in single turnover experiments with DNA templates showed that pyrophosphate (PPi) dissociation was fast following nucleotide incorporation so that it did not contribute to enzyme specificity ($k_{cat}/K_m$) (Chapter 2). Interestingly, for the first time it was shown that PPi release was rate-limiting when a RNA template was used (Chapter 4). Analysis is presented to show how the slow PPi dissociation alters nucleotide selectivity by HIV-1 RT (Chapter 5).

**Coordination Between the Polymerase and RNase H Active Sites**

There has been considerable debate over the presence of a temporal and spatial coordination between the two distal active sites of HIV-1 RT, the polymerase and RNase H active sites. Several biochemical data have supported the existence of some coordination. First, RNA cleavage was observed 18 or 19 nucleotides upstream from the 3'-OH of the growing DNA primer in the presence of trap that prohibited substrate rebinding, indicating the extension and degradation activities were able to be achieved by a single binding event (*92*). A more definitive study showed that DNA polymerization

and RNase H hydrolysis occur coincidently on the ms time scale with 18 base-pairs between the two sites (*20*). Then, a more recent biochemical study suggested that HIVRT can simultaneously engage its DNA/RNA substrate at both active sites based on the observations that small ligands capable of stabilizing the nucleic acid in the polymerase active site only affected the pattern but not the efficiency of RNA cleavage under single-turnover conditions (*93*). However, structural analyses have suggested that a DNA/RNA hybrid is not able to engage the polymerase and RNase H active sites of HIV-1 RT simultaneously (*15, 51, 54*). Alternatively, these structural analyses suggested a 'toggling' model, where the bound substrate must toggle between the two active sites of HIV-1 RT. The 'toggling' model was supported by insufficient biochemical analyses. First, the rate of primer extension was determined to be 7-fold faster than that of RNA cleavage (*20*) and the template was not degraded processively in RNA dependent extension assays (*19*). Moreover, RNA cleavage tended to occur more efficiently when primer extension paused before a hairpin on the RNA template (*94*) and NNRTIs inhibitory to DNA synthesis were able to activate RNA cleavage (*95*). Therefore, whether the two catalytic centers coordinate with each other during the reverse transcription is still elusive and will be further explored in this dissertation.

**Single Molecule Studies**

HIV-1 RT is multi-function enzyme that participates in DNA- and RNA-dependent DNA synthesis (including strand-transfer and strand-displacement synthesis), as well as DNA-directed RNA hydrolysis. Crystallographic and ensemble kinetic studies have demonstrated certain conformational dynamics of HIV-1 RT during different stages

of DNA synthesis. However, how RT recognizes different substrates and performs corresponding functions is still unknown. Single-molecule FRET (smFRET) assays revealed that RT bound to DNA and RNA primers with opposite orientations, directing primer extension or template degradation activities. When a PPT primer was utilized for (+) strand DNA synthesis, HIV-1 RT could flip between the two opposite orientations and the binding of next correct nucleotide stabilized the polymerization-competent mode (*96*). In addition, a similar smFRET study showed that RT did not directly bind to the 3' terminus of the growing primer. Alternatively, it slided a long range on the nucleic acid substrate and found the DNA 3' end. The sliding and flipping mechanisms may facilitate multiple stages of the reverse transcription pathway, especially for the strand-transfer and strand-displacement synthesis (*97*). Ensemble kinetic studies of DNA synthesis on the hairpin containing templates revealed the productive and nonproductive binding modes of RT to the substrates. If RT binds in a nonproductive mode, it can switch to the productive mode by melting the next stem base-pair (*98-99*). Single-molecule DNA flow-stretching assay was performed to address the question how RT replicates through secondary structures on RNA template or nacent single-stranded DNA (ssDNA) without the help from a motor protein (e.g., a helicase) for unwinding. Their results revealed that the strand-displacement synthesis by RT utilized the combination of active (dNTP incorporation) and passive (thermal fraying) mechanisms (*100*). However, these studies merely reproduced what had been defined by more comprehensive and quantitative ensemble studies a decade earlier (*98, 101-102*). Single molecule studies also provided kinetic insights on how HIV-1 RT initiates the (-) strand DNA synthesis. *In vitro* studies

showed that the addition of first few nucleotides was slow during the initiation process, but then speeded up after the successful insertion of 5-6 nucleotides (*103-104*). Single-molecule FRET (smFRET) assays revealed the mechanistic basis for this phenomenon by demonstrating that HIV-1 RT transitions between two opposite binding orientations and pauses are corresponding to the polymerization-incompetent configuration. Destabilization of the stem-loop structure within viral RNA triggers the initiation-to-elongation transition (*105*).

In summary, biochemical and crystallographic studies have led to a better understanding of the nucleotide selectivity and drug resistance by HIV-1 Reverse Transcriptase in the past 20 years. However, questions remain in this field due to inadequate experimental design and flawed kinetic analysis. In addition, the role of pyrophosphate release in nucleotide specificity by HIV-1 RT has been overlooked. On the other hand, the limited number of structural analyses failed to demonstrate the potential coordination between the polymerase and RNase H active sites of HIV-1 RT. These questions, however, are explored in this dissertation by pre-steady-state kinetics and global data analysis. Data presented in the following chapters provide insights into HIV-1 RT nucleotide selectivity, drug resistance and RNase H activities.

# Chapter 2: Mechanistic Basis for AZT Drug Resistance Conferred by Dipeptide Insertions in the Fingers of HIV-1 Reverse Transcriptase

## 2.1 INTRODUCTION

HIV-1 reverse transcriptase (HIVRT) is a heterodimer, consisting of two related subunits: p66 contains both polymerase and RNase H sites, while p51 is derived from p66 by proteolytic removal of the C-terminal RNase H domain (*30*). The RNA- and DNA-dependent polymerase and the RNase-H activities are essential for converting the viral RNA genome into dsDNA that can be integrated into the host genome (*106-107*). Due to its vital role at the beginning of the viral life cycle, RT has been an important target for antiretroviral therapy (*60-61*). Two distinct groups of compounds were developed targeting HIVRT: nucleoside/nucleotide analog reverse transcriptase inhibitors (NRTIs) and non-nucleoside reverse transcriptase inhibitors (NNRTIs). NRTIs are a group of structurally diverse analogues of natural nucleoside/nucleotides, which lack the 3'-hydroxyl group and therefore can serve as chain terminators after incorporation. Among these compounds, AZT was the first approved by the US Food and Drug Administration (FDA) in 1987 and has been widely used in the clinic until now.

Sustained viral suppression requires long term treatment (*66*), which often results in the appearance of viruses resistant to these compounds. Mutational patterns in the RT gene associated with drug resistance to every NRTI have been identified and some complexes of mutations were found to confer cross-resistance to all (or almost all) NRTIs currently approved by the US FDA, such as the Q151M complex, thymidine analog mutations (TAMs) and the 69 insertion complex (*67*). TAMs mutations (M41L, D67N,

K70R, L210W, T215F/Y, and K219Q/E) appeared following prolonged treatment with AZT and are associated with decreased susceptibility of HIV to AZT (*108*), as well as all other NRTIs (*67*). Dipeptide insertions between residues 69 and 70 of HIVRT (most commonly Ser-Ser/Ser-Gly) were observed in clinic samples of patients receiving combination therapy with AZT and other drugs (ddC or ddI) (*109-110*). These insertions generally occurred in the background of AZT resistance mutations (TAMs) and were therefore termed the 69 insertion complex (*111-113*). Among TAMs mutations, D67N was rarely found in 69 insertion complex mutations (*114-115*) but A62V, which is not a common mutation for AZT resistance, was frequently found in this complex. Structurally, A62V was reported to interact with M41L and thereby affect the positioning of the β3-β4 loop (*116*). Fingers insertion in the presence of TAMs have been reported to enhance ATP-dependent excision of thymidine analogues, including AZT (*117-119*), as well as cause lower susceptibility to almost all other NRTIs (*109, 117, 120-121*).

Two primary mechanisms have been proposed for RT resistance to NRTIs: i) enhanced discrimination against NRTI incorporation, and; ii) enhanced removal of chain terminators from the 3'-end of the blocked primer, so DNA synthesis can resume (*69*). In nucleotide-dependent excision, ATP reacts with the 3' terminal base to produce a dinucleotide tetraphosphate and a 3'-OH terminated primer capable of subsequent extension. TAMs are a well studied example for the excision mechanism (*79, 122*), which has been proposed to facilitate excision of thymidine analogues either by promoting a proper orientation of the pyrophosphate donor molecule (*115*) or by creating a high-affinity ATP-binding site (*76*). However, comparable dissociation constants of ATP were

obtained from kinetic analysis for WT and AZT resistant mutants (*78*). A recent single turnover kinetic analysis showed that the minimal 2-fold discrimination against AZT relative to TTP by TAMs was due to the fact that the slower rate of AZT incorporation by TAMs was offset by a lower $K_m$, resulting from the equilibration of the conformational change step (*77*). However, the mechanism by which the 69 insertion complex enhances the resistance to AZT and other NRTIs is still controversial. Some studies showed that mutants with fingers insertions can interfere with incorporation of certain drugs, especially, 3TCTP (*117*) while others suggested that these mutations showed no or only limited discrimination against incorporation of chain terminators (*119, 123*). Several reports have indicated that the resistance phenotype associated with the fingers insertion complex involves ATP-mediated excision (*115-117, 119-120, 124*), and the role of the dipeptide insertion is to destabilize the formation of ternary dead end complex (DEC) (*117*) and thereby reduce the inhibition of ATP-dependent unblocking reaction by the next complementary dNTP (*119-120*). However, some of the published excision data are not so convincing because the data were obtained either at a fixed time point under various ATP concentrations or at different time points but with fixed ATP concentration. Consequently, key information about excision rates ($k_{exc}$), ATP binding affinity ($K_{m,ATP}$), as well as excision efficiency ($k_{exc}/K_{m,ATP}$) was unfortunately missed. Meanwhile, some other data are hard to interpret because the protein used in these studies were derived from clinical isolates harboring complex suites of mutations in addition to fingers insertions and AZT resistant mutations.

Whether fingers insertion has any effect on discrimination against incorporation of NRTIs is not well determined so we have revisited this problem using AZT as an example. To address this question, pre-steady-state kinetic analysis was performed with recombinant proteins harboring simple set of mutations (T69SSS in the otherwise WT RT or in the TAMs mutant). Specificity constants of TTP and AZTTP were compared to quantify and examine the mechanistic basis for changes in AZT discrimination afforded by the dipeptide insertion. Also, primer unblocking and extension assays were performed by analysis of the concentration dependence and time course of reaction to assess the role of fingers insertion on excision efficiencies and measure the binding affinity of the next cognate nucleotide.

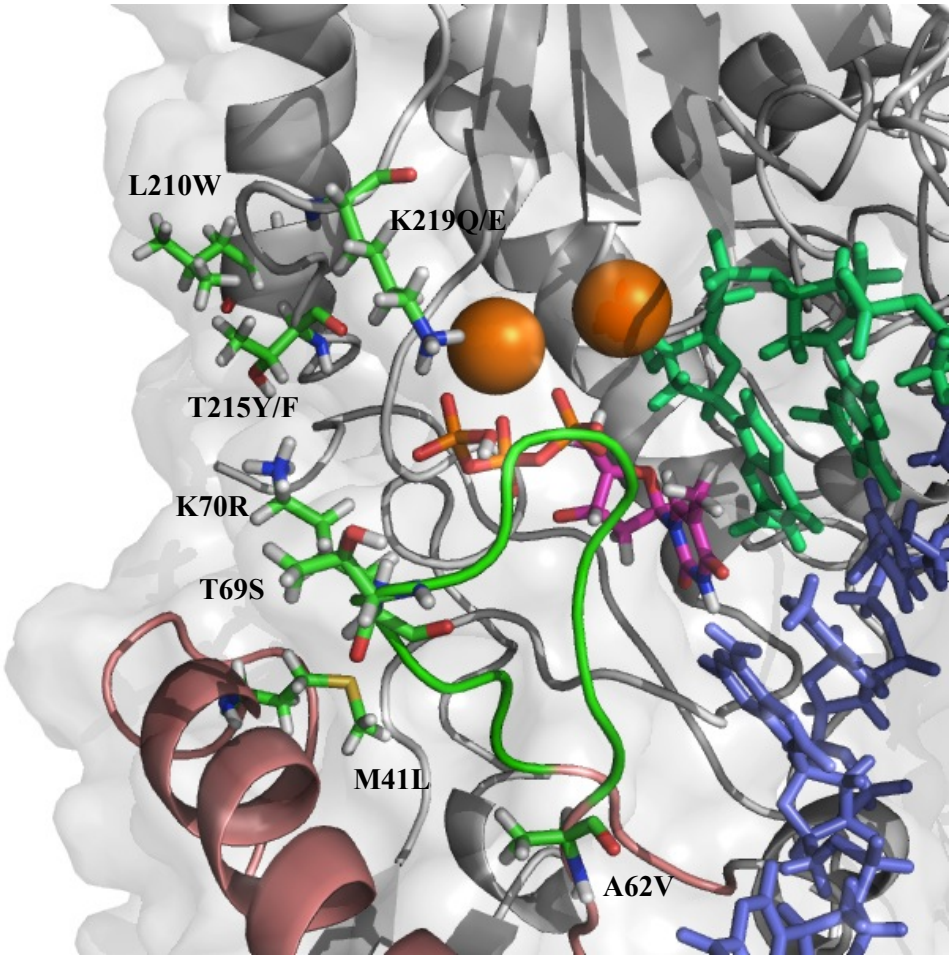



**Figure 2.1** *69 Insertion complex mutations of HIV RT.* The crystal structure of HIV RT (PDB: 1RTD) in complex with a DNA primer/template (green and blue) and TTP (color coded by element). The 69 insertion complex has the TAMs backbone (M41L, K70R, L210W, T215F/Y, and K219Q/E) , together with A62V, T69S, and a dipeptide insertion between residues 69 and 70 of HIVRT (most commonly Ser-Ser/Ser-Gly, not shown in the figure). Among TAMs mutations, D67N was rarely found in 69 insertion complex mutations, but A62V, which is not a common mutation for AZT resistance, was frequently found in this complex and reported to interact with M41L and thereby affect the positioning of the β3-β4 loop (*116*).

## 2.2 MATERIALS AND METHODS

### Expression and Purification of MDCC-labeled HIV Reverse Transcriptase

The T69SSS and TAMs_T69SSSVD HIV-1 RTs were expressed, purified and labeled following methods previously described (*77, 86*). Briefly, the T69SSS mutant was produced by introducing the T69S substitution and a two-serine insertion between position 69 and 70 (T69SSS) in genes of both subunits using a Quikchange Multi kit (Stratagene). Based on genes of the thymidine analogues mutant (TAMs: M41L, D67N, K70R, L210W, T215Y and k219E), the TAMs_T69SSSVD mutant was produced by introducing T69SSS followed by the introduction of A62V and a reverse mutation of N67D (Table 2.1). Two subunits for each mutant were individually expressed in T7 Express Competent *E.coli* (New England Biolabs; NEB), mixed in a 1:1 ratio and purified by tandem Q-Sepharose and Bio-Rex 70 columns followed by a single-strand DNA affinity column. Following MDCC labeling, excess MDCC was removed using a Bio-Rex 70 column. The labeled enzymes were assayed by pre-steady-state burst experiments to determine the active concentrations, divided into aliquots, and stored at -80°C.

**Table 2.1** *Mutagenic T69SSS and TAMs_T69SSSVD primers*

PRIMER 1: T69S-SS

5′-TATTTGCCATAAAGAAAAAAGACAGT**TCTTCTTCT**AAATGGAGAAAATTAGTAGATTTCA-3′

PRIMER 2: TAMS-T69S-SS

5′-TATTTGCCATAAAGAAAAAAAAACAGT**TCTTCTTCT**CGCTGGAGAAAATTAGTAGATTTCA-3′

PRIMER 3: TAMS-T69S-SS-A62V-N67D

5′-CATACAATACTCCAGTATTT**GTC**ATAAAGAAAAAA**GAC**AGTTCTTCTTCTCGCTGGAGAA-3′

Mutated codons are shown in bold and underlined. Primers 1 and 2 were used to introduce T69S-SS in WT or TAMs backbone. Primer 3 was used to introduce A62V and N67D into TAMs-T69S-SS.

**DNA Substrates for Kinetic Studies**

The 25mer and 45mer DNA ordered from Integrated DNA Technologies (Table 2.2) were purified by gel extraction individually. Annealing of the 25mer primer and 45mer template was carried out by mixing the two oligonucleotides in a 1:1 molar ratio at 95°C for 5min, followed by slow cooling to room temperature. The primer was 5' -$^{32}$P -labeled with T4 polynucleotide kinase (NEB) for use in the rapid quench flow kinetic assays. Dideoxynucleotide-terminated oligo-primer used in nucleotide off-rate assay (25ddA) was made through enzymatic synthesis with RT, followed by gel-extraction purification.

**Table 2.2** *DNA substrates for kinetic studies*

*25/45-A:*

25mer: 5′-GCCTCGCAGCCGTCCAACCAACTCA-3′

45mer: 3′-CGGAGCGTCGGCAGGTTGGTTGAGT**A**GCAGCTAGGTTACGGCAGG-5′

*25ddA/45-A:*

25mer: 5′-GCCTCGCAGCCGTCCAACCAACTCA$_{dd}$-3′

45mer: 3′-CGGAGCGTCGGCAGGTTGGTTGAGT**A**GCAGCTAGGTTACGGCAGG-5′

**Quench Flow Kinetic Assays**

The time dependence of nucleotide incorporation was measured by single-turnover rapid quench flow techniques as described (*77*) using a KinTek RQF-3 (KinTek Corp.). A preformed enzyme-DNA duplex (150 nM of MDCC-labeled HIV RT and 100 nM 25/45 DNA) was rapidly mixed with various concentrations of nucleotide using the RQF-3 Quench Flow (KinTek). All reactions were conducted in 50 mM Tris pH7.5, 100 mM KCl and 10 mM magnesium acetate (all concentrations final) at 37°C and the reaction was stopped by the addition of 0.5 M EDTA. Products were separated by 15% denaturing PAGE and quantified using a Typhoon scanner in ImageQuant 6.0 (GE). Primer block excision and extension was measured by a hand-quench assay. The 25/45mer template-primer (100 nM) was pre-incubated with either WT or mutant RTs (150 nM) at 4°C for 10 min in the same buffer for nucleotide incorporation assays. The primer was then extended by adding equal volume of AZTTP (12.5 μM) and incubating at 37°C for 5 min to form an AZT-terminated primer. Primer rescue and extension was initiated by adding 100 μM TTP and 1 μM dCTP in the presence of various concentrations of ATP and 0.6 μM pyrophosphatase. A 10 μL aliquot of the reaction was

mixed with 10 µL 0.5 M EDTA and 20 µL loading buffer to stop the reaction for each time point. Products were then separated and quantified as described above. The experimental setup was similar in assays designed to assess the inhibitory effect of dCTP (the next correct nucleotide at the 3'-end of AZT-terminated primer) except that the concentration of ATP was fixed at 4.8 mM while various concentrations of dCTP were used.

**Stopped Flow Kinetic Assays**

The time dependence of transient MDCC fluorescence changes corresponding to nucleotide induced enzyme conformational change was monitored using an Auto SF-200x (KinTek Corp.). A pre-formed enzyme-DNA duplex (100 nM of MDCC-labeled HIV RT and 150 nM 25/45 DNA) was rapidly mixed with various concentrations of nucleotide in the stopped flow for 300-500 ms. MDCC fluorescence was excited at 425 nm and fluorescence was monitored at 475 nm with a 50 nm band width band-pass filter (Semrock). To determine the nucleotide dissociation rate, 200 nM enzyme-$DNA_{dd}$ (2', 3'-dideoxy-terminated primer) was pre-incubated with 2.5 to 10 µM of nucleotides (TTP or AZTTP), depending on their binding affinities. The preformed ternary complex was then mixed with unlabeled enzyme-DNA duplex to serve as a trap for free nucleotide. A coupled fluorescence assay described elsewhere (*125*) was used to measure the kinetics of pyrophosphate (PPi) release. Briefly, the preformed enzyme-DNA duplex (150 nM of HIV RT and 100 nM 25/45 DNA) was rapidly mixed with 25 µM of nucleotide in the presence of 0.6 µM of pyrophosphatase (PPase), the Pi "mop" containing 100 µM of 7-

methylguanosine (7-MEG) and 0.02 IU/mL purine nucleoside phosphorylase (PNPase), as well as 1.5 µM of fluorescently labeled *E.coli* phosphate binding protein (PBP).

**Global Analysis of Kinetic Data**

The kinetic constants governing nucleotide binding and incorporation were obtained by a global fitting of the data to a two-step nucleotide binding mechanism (Scheme 2.1) using KinTek Explorer software (KinTek Corp.) Rate constants governing DNA binding ($k_{on}$) and dissociation ($k_{off}$) were fixed based on previous estimates while those governing the isomerization step ($k_2$) were fixed based on estimates from temperature dependent assays.

**Scheme 2.1** *Two-step Nucleotide Binding Induced-fit Mechanism*

$$\overset{Open}{ED_n} + dNTP \underset{}{\overset{K_1}{\rightleftharpoons}} \overset{Open}{ED_n dNTP} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \overset{Closed}{FD_n dNTP} \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} \overset{Closed}{FD_{n+1} PPi} \xrightarrow{fast} \overset{Open}{ED_{n+1}} + PPi$$

$$E + D_n \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} ED_n \; ; \; E + D_{n+1} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} ED_{n+1}$$

Nucleotide dissociation rates were measured by a trap assay and analyzed using a previously developed mechanism (Scheme 2.2) (*77*).

**Scheme 2.2** *Measuring the Nucleotide Dissociation Rate by Competition*

$$\overset{Open}{ED_{dd}} + dNTP \underset{}{\overset{K_1}{\rightleftharpoons}} \overset{Open}{ED_{dd} dNTP} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \overset{Closed}{FD_{dd} dNTP} \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} \overset{Closed}{GD_{dd} dNTP}$$

$$dNTP + Trap \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} dNTP.Trap$$

Kinetics of PPi release were also included in the global fitting (Scheme 2.3) and kinetics of pyrophosphatase, as well as phosphate binding to MDCC-PBP were determined to calibrate the kinetics of the sensor. F represents MDCC-PBP, F~P represents the phosphate induced high fluorescent state of MDCC-PBP, and G represents pyrophosphatase (PPase). Pyrophosphate (PP) released from the reaction is captured by the pyrophosphatase ($k_1$) and undergoes two steps of hydrolysis ($k_2$ and $k_3$) to release two inorganic phosphates (P). In order to reduce the number of variables, pyrophosphate binding and hydrolysis steps are all set to be irreversible. The released phosphates then bind to the MDCC-labeled PBP ($K_1$), which isomerizes ($K_2$) to the high-fluorescent state (F~P) and yields the signal we measured.

**Scheme 2.3** *Calibraiton of the kinetics of the fluoresent sensor (MDCC-PBP)*

$$F + P \underset{}{\overset{K_1}{\rightleftharpoons}} FP \overset{K_2}{\rightleftharpoons} F \sim P$$

$$G + PP \xrightarrow{k_1} G.PP \xrightarrow{k_2} GP + P;\ GP \xrightarrow{k_3} G + P$$

Excision efficiencies were determined by ATP concentration dependent rescue assays while the inhibitory effect of dCTP was examined through similar assays with fixed ATP concentration and various nucleotide concentrations. The data were globally fit to a model including a translocation step between the nucleotide binding site (N site) and the priming site (P site) (*79*), as shown in Scheme 4.

**Scheme 2.4** *ATP Mediated Excion of AZTMP or Formation of Dead End Complex (DEC)*

$$ED_n^{P\ site}AZTMP + dNTP \underset{}{\overset{K_{d,1}}{\rightleftharpoons}} ED_n^{N\ site}AZTMP.dNTP$$

$$K_0 \Updownarrow$$

$$XD_n^{N\ site}AZTMP + ATP \underset{}{\overset{K_m}{\rightleftharpoons}} XD_nAZTMP.ATP \underset{k_{-2}}{\overset{k_{exc}}{\rightleftharpoons}} ED_n^{P\ site} + AP_4AZT$$

## 2.3 RESULTS

### TTP and AZTTP Incorporation by the HIV RT_T69SSS-MDCC

In order to identify the rate constants governing nucleotide binding and incorporation, four sets of experiments (Figure 2.2 and 2.5) were conducted and fit simultaneously to a single model as illustrated by Scheme 2.1, 2.2 and 2.3. Figure 2.2A shows the time course of fluorescence changes upon nucleotide binding and incorporation at various nucleotide concentrations. Figure 2.2B shows time dependence of nucleotide incorporation measured by rapid chemical quench flow method at several nucleotide concentrations. Figure 2.2C shows the time course of fluorescence increase occurring when nucleotide dissociates from the enzyme-$DNA_{dd}$-TTP ternary complex and is then trapped by unlabeled enzyme-DNA duplex. Figure 2.2D shows the time dependence of the fluorescence change measured using the stopped flow fluorescence coupled assay to monitor pyrophosphate (PPi) release, hydrolysis to Pi and then binding of Pi to the fluorescent Pi binding protein sensor. The data were fit globally, as illustrated by the smooth curves in each figure panel, to derive the rate constants summarized in Table 2.3.

**Table 2.3** *Kinetic Constants Governing Nucleotide Binding and Incorporation*

| Enzyme | dNTP | $1/K_1$ (μM) | $k_2$ ($s^{-1}$) | $k_{-2}$ ($s^{-1}$) | $K_2$ | $k_3$ ($s^{-1}$) |
|---|---|---|---|---|---|---|
| T69SSS | TTP | 274±2 | 2140±180 | 1.5±0.1 | 1430±150 | 21.1±0.1 |
| | AZTTP | 25±2 | 590±30 | 31.8±1.4 | 18.6±1.3 | 15.8±0.1 |
| TAMs_T69SSSVD | TTP | 2840±12 | 4350±380 | 1.5±0.1 | 2900±320 | 11.6±0.1 |
| | AZTTP | 26±2 | 105±11 | 48.4±3.7 | 2.2±0.3 | 11.0±0.6 |

Rate constants were derived from global analysis of the data fitted to Scheme 2.1 allowing for a nucleotide induced conformational change prior to chemistry. Values for $k_2$ were determined by temperature dependence analysis at lower temperatures and the rate at 37°C was determined by extrapolation based on Arrhenius analysis to constrain the global fit. Global analysis of $K_1$ was performed by locking the initial second-order nucleotide on-rate ($k_1$) to a modest estimate for diffusion-limited collision (100 $\mu M^{-1}s^{-1}$). The reverse rate ($k_{-1}$) was fit globally. The reported ground-state $K_d$ ($1/K_1$) and equilibrium constant of conformational change step ($K_2$) were calculated using the values obtained from global analysis.

**Table 2.4** *Nucleotide Binding, Specificity, and Discrimination for WT_T69SSS and TAMs _T69SSSVD*

| Enzyme | dNTP | $K_{d,net}$ (μM)[a] | $K_m$ (μM) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($\mu M^{-1}s^{-1}$)[b] | D[c] |
|---|---|---|---|---|---|---|
| T69SSS | TTP | 0.20±0.02 | 3.1±0.3 | 20.9±2.5 | 6.8±0.6 | |
| | AZTTP | 1.3±0.1 | 2.0±0.2 | 14.6±1.1 | 7.3±0.7 | 0.93±0.11 |
| TAMs_T69SSSVD | TTP | 1.0±0.1 | 8.6±0.8 | 11.7±1.4 | 1.4±0.1 | |
| | AZTTP | 11.8±1.9 | 9.5±1.2 | 7.0±0.9 | 0.7±0.1 | 2.0±0.3 |

[a]The net $K_d$ as a result of two-step nucleotide binding was calculated as $1/(K_1(1+K_2))$.
[b]$k_{cat}/K_m$ was calculated as $k_1k_2k_3/(k_2k_3+(k_{-1}(k_{-2}+k_3)))$
[c]Discrimination was calculated as $D = (k_{cat}/K_m)_{TTP}/(k_{cat}/K_m)_{AZTT}$

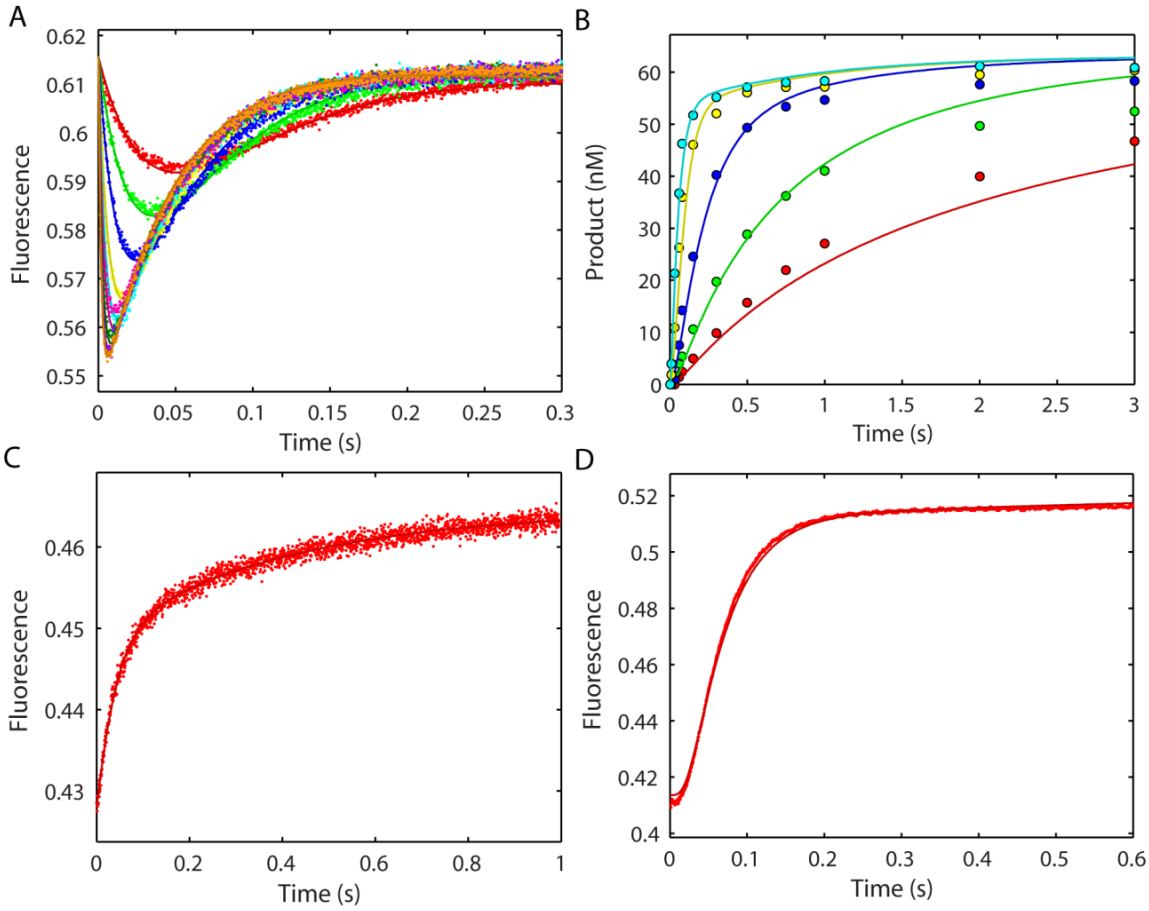



**Figure 2.2** *Global fitting of TTP Binding and Incorporation by MDCC-labeled HIV Reverse Transcriptase T69SSS.* (A) TTP concentration dependence of MDCC fluorescence to monitor nucleotide binding and incorporation. Various concentrations of TTP (2.5, 5, 10, 20, 30, 40, 60, 80, and 100 μM) were mixed with a pre-formed enzyme-DNA complex (150 nM 25/45mer and 100 nM MDCC-labeled HIV RT_T69SSS). (B) Rapid quench flow TTP concentration dependence of the rate of product formation. Various concentrations of TTP (0.1, 0.25, 0.75, 3, 15 μM) were mixed with a pre-formed enzyme-DNA complex (75 nM 25/45mer DNA and 175 nM MDCC-labeled HIV RT_T69SSS). (C) The rate of re-opening of the 'closed' ternary complex was measured by rapidly mixing a pre-formed enzyme-$DNA_{dd}$-TTP complex (200 nM HIV RT_T69SSS-MDCC, 300 nM 25ddA/45mer, 2.5 μM nucleotide) with a nucleotide trap consisting of 2.5 μM unlabeled enzyme-DNA complex. (D) Rate of PPi release during the incorporation of TTP was monitored by rapidly mixing 25 μM of TTP with a pre-formed enzyme-DNA complex (100nM 25/45mer DNA and 150 nM MDCC-labeled HIV RT_T69SSS) in the presence of 1.5 μM of MDCC-PBP and 0.6 μM of PPase. In each panel the smooth lines show the best fit achieved by globally fitting data in panels A-D simultaneously. Rate constants are summarized in Table 2.3.

Because the maximum rate of conformational change ($k_2$) was too fast to measure at 37°C (Figure 2.2A), the experiments were repeated at lower temperatures (5°C, 10°C, 18°C and 25°C) in order to estimate the maximum rate of nucleotide-induced conformational change ($k_2$) by extrapolation to 37°C. Analysis of the TTP concentration dependence of the rate of the decrease of fluorescence gave estimates of $k_2$ = 104±6, 178±4, 353±17, and 749±18 $s^{-1}$; as well as estimates of $K_d$ = 41±5, 59±3, 126±7, and 154±7 μM at 5, 10, 18, and 25°C, respectively (Figure 2.3A/B). A maximum rate of 2140±180 $s^{-1}$ at 37°C was obtained by extrapolation (Figure 2.3C). This is in agreement with the lower and upper limits of $k_2$ (650~5000 $s^{-1}$) estimated by confidence contour analysis while globally fitting the data shown in Figure 2.2. The $K_d$ values can also be used in the Van't Hoff analysis to derive thermodynamic parameters governing nucleotide binding, in particular the enthalpy change (Figure 2.3D). To calculate the free energy, a standard concentration of nucleotide substrate was set as 100 μM, which is the approximate physiological concentrations of nucleotides in living cells. The data showed the free energy of ground state TTP binding is favorable at lower temperatures (5°C and 10°C, where $K_d$ < 100 μM) but unfavorable at higher temperatures (18°C and 25°C, where $K_d$ > 100 μM). The value of $k_2$ at 37°C obtained from Arrhenius analysis was fixed and all other rate constants were then derived from the global data fitting. The second-order rate constant of TTP binding $K_1k_2$ = 7.1±0.3 $\mu M^{-1}s^{-1}$ was accurately determined from the data obtained at 37°C (dash line, Figure 2.3A). Since $k_2$ = 2140 $s^{-1}$, an estimate of $1/K_1$ = 300±28 μM can be calculated. This value is comparable to the

ground state dissociation constant of $1/K_1 = 274\pm2$ μM that was obtained from the global data fitting. Therefore, results obtained from conventional data analysis provide an additional check of the model used to fit data by global analysis.

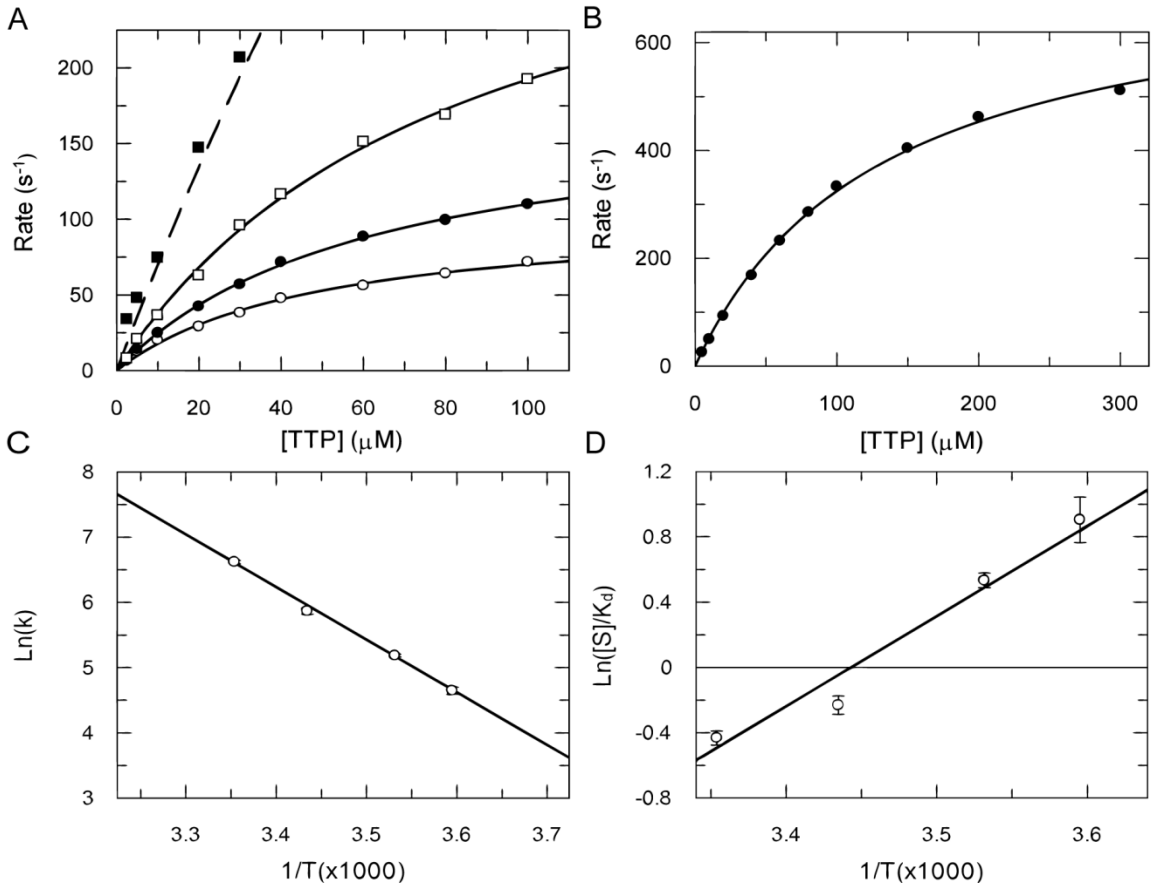



**Figure 2.3**

**Figure 2.3** *Temperature Dependence of the HIV RT_T69SSS Conformational Change Rate upon TTP Binding*. (A) The experiment shown in Figure 2.2A was repeated at various temperatures. A pre-incubated complex of 150 nM 25/45mer and 100 nM MDCC-labeled HIV RT_T69SSS was rapidly mixed with various concentrations of TTP (2.5, 5, 10, 20, 30, 40, 60, 80, and 100 μM) at 5°C (○), 10°C (●), 18°C (□) and 37°C (■). (B) A pre-incubated complex of 150 nM 25/45mer and 100 nM MDCC-labeled HIV RT_T69SSS was rapidly mixed with various concentrations of TTP (5, 10, 20, 40, 60, 80, 100, 150, 200 and 300 μM) at 25°C. The concentration dependence of the fast decrease of fluorescence (corresponding to conformational closing) was fit to hyperbolic equation to obtain the maximum rate of the conformational change ($k_2$) at each temperature resulting in rates of 104±6, 178±4, 353±17, and 749±18 $s^{-1}$; as well as estimates of $K_d$ = 41±5, 59±3, 126±7, and 154±7 μM at 5, 10, 18, and 25°C, respectively. The data corresponding to the concentration dependence of the conformational change rate at 37°C were fit to a linear equation because rates at higher concentrations of nucleotide were not accurately determined due to the dead time limitation of the instrument and the fast rate of the reaction. (C) The temperature dependence of $k_2$ was analyzed by Arrhenius plot to estimate a maximum rate for $k_2$ of 2140±180 $s^{-1}$ at 37°C. This value was then fixed to constrain data fitting by global analysis. (D) The temperature dependence of $K_d$ was analyzed by Van't Hoff plot to derive estimates of ΔH of -48.4± 6.3 kJ/mol, ΔS of 0.17±0.02 kJ $K^{-1}$/mol, as well as ΔG of -2.1±0.3, -1.3±0.1, 0.56±0.1, and 1.1±0.1 kJ/mol at 5, 10, 18, and 25°C, respectively.

Next, the rate of TTP release from an E-$DNA_{dd}$-TTP ternary complex was measured by mixing the complex with unlabeled enzyme-DNA duplex to trap free TTP. 3'-dideoxy-terminated primer was used in this assay to block the chemistry step. Since the ground state binding of TTP is a fast equilibrium process, the reverse of the conformational change is the limiting step for TTP dissociation and was actually determined by the assay. This yielded a $k_{-2}$ value of 1.5±0.1 $s^{-1}$, which was well constrained in the global fitting according to confidence contour analysis. It is worth noting that a three-step binding mechanism was used to fit the biphasic data obtained from the competition assay. Specifically, an additional isomerization step (Scheme 2.2) was included to account for the slower phase.

The rate of PPi release was examined using a coupled assay previously developed in our lab (*125*). Briefly, PPi released from the reaction is hydrolyzed by the pyrophosphatase (PPase), yielding inorganic phosphates, which then bind to the fluorescently labeled *E.coli* phosphate binding protein (PBP) and give the fluorescence signal. The kinetics of PPi hydrolysis and Pi binding to MDCC-PBP are calibrated by two sets of experiments (Figure 2.4). Binding of Pi to PBP includes two steps: a weak ground state binding ($K_1$ =10 μM) followed by a rapid isomerization to the highly fluorescent state (Scheme 2.3). The rate constants obtained by fitting these two experiments were then used in the global fitting to derive the rate of PPi release. This afforded a PPi release rate of 2360±680 $s^{-1}$, which unfortunately was not well constrained since the rate was too fast to be accurately determined (Figure 2.2D). The confidence contour analysis set a lower limit for the rate as 203 $s^{-1}$ (i.e., it must be ~10X faster than

catalysis); therefore the rate was fixed at 1000 $s^{-1}$ for simplicity in the global fitting to derive the remaining rate constants.

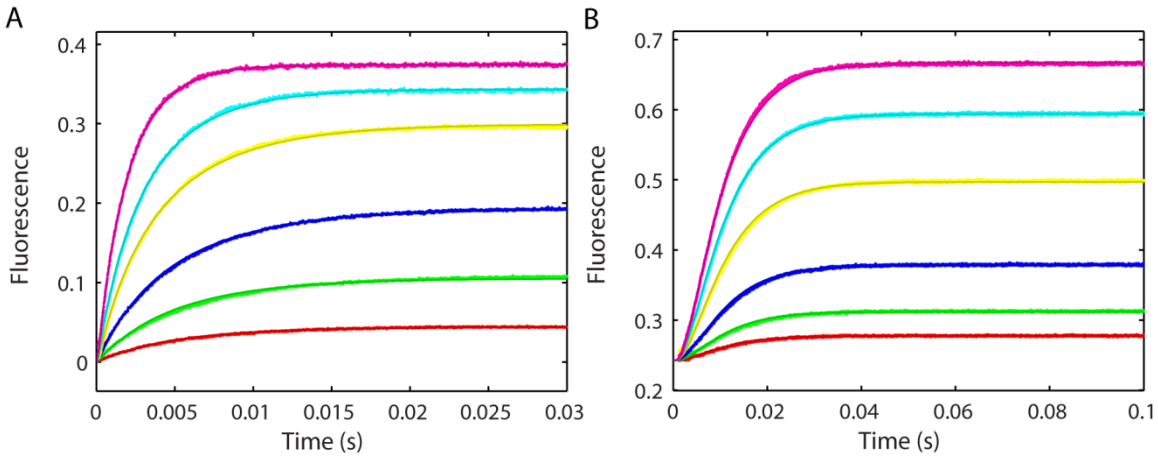



**Figure 2.4** *Calibration of the kinetics of the sensor.* (A) Kinetics of Pi binding to fluorescent phosphate sensor (MDCC-PBP). Various concentrations of Pi (0.2, 0.5, 1, 2, 3, and 5 μM) were rapidly mixed with 1.5 μM of MDCC-PBP in the presence of Pi "mop", using a stopped flow instrument. (B) Kinetics of pyrophosphatase (PPase) measured using the fluorescent phosphate sensor. Various concentrations of PPi (0.05, 0.1, 0.2, 0.4, 0.6, and 0.8 μM) were rapidly mixed with 1.5 μM of MDCC-PBP in the presence of 0.6 μM of PPase and the Pi "mop". Panel A and B were fit globally according to scheme 2.3. Rate constants obtained from global analysis were then used to derive the rate of PPi release in subsequent experiments.

Similar experiments and data analysis were performed to determine kinetics governing AZTTP binding and incorporation as shown in Figure 2.5 and summarized in Table 2.3. The ground state dissociation constant of AZTTP binding ($1/K_1$) is 10-fold lower than that of TTP binding, with value of 25±2 μM, indicating a much tighter binding of AZTTP to the active site. In the case of AZTTP incorporation, the rate of conformational change decreased to a measurable range (<600 $s^{-1}$), so Arrhenius analysis was not performed and a rate of 590±30 $s^{-1}$ was derived from the global fitting. The

nucleotide dissociation-rate was fit to a rate of 31.8±1.4 $s^{-1}$. The rate of chemistry was 15.8±0.1 $s^{-1}$, which is a little bit slower than for of TTP incorporation. This rate was defined by the slow second phase of the stopped flow data (Figure 2.5A), the rapid chemical quench flow data (Figure 2.5B), as well as the PPi release data (Figure 2.5D). Again, PPi release was fast, with a rate of 520±65 $s^{-1}$ (Figure 2.5D) and a lower limit of that rate as 213 $s^{-1}$, determined by confidence contour analysis.

**TTP and AZTTP Incorporation by MDCC labeled HIVRT TAMs_T69SSSVD**

Four experiments were performed to determine the rate constants governing the binding and incorporation of TTP by the TAMs_T69SSSVD mutant. The concentration dependence of nucleotide incorporation in the chemical quench flow data and that of fluorescence change in the stopped flow data determined the ground state dissociation constant as 2840±12 μM (Figure 2.6A/B). Temperature dependence of the conformational change rate was studied, yielding rates of 205±3, 332±9 and 770±29 $s^{-1}$; as well as estimates of $K_d$ = 402±2, 600±2, and 980±2 μM for 5, 10, and 18°C (Figure 2.7A). The temperature dependence of the rates was analyzed on an Arrhenius plot (Figure 2.7B) and extrapolation afforded an estimate of a maximum rate of $k_2$ of 4350±380 $s^{-1}$ at 37°C. Since $K_1k_2$ = 1.54±0.05 $\mu M^{-1}s^{-1}$ was accurately determined, an estimate of $1/K_1$ = 2820±260 μM can be calculated. Again, this value is consistent with the ground state dissociation constant of $1/K_1$ = 2840±12 μM that was obtained from the global data fitting. Thermodynamic constants were also derived by Van't Hoff analysis (Figure 2.7C), yielding free energy change values (ΔG) of 3.2±0.1, 4.2±0.1, and 5.5±0.1

KJ/mol for 5, 10, and 18°C. So the binding of TTP to TAMs_T69SSSVD mutant was unfavorable at all temperatures tested. The dissociation rate of TTP ($k_{-2}$) was 1.5±0.1 $s^{-1}$, as measured by the nucleotide off-rate assay (Figure 2.6C). The rate of chemistry is defined by three experiments (Figure 2.6A/B/D), and therefore was well constrained to a value of 11.6±0.1 $s^{-1}$. Measurement of the PPi release was fit to a rate of 340±23 $s^{-1}$ (Figure 2.6D), with its lower limit of 90 $s^{-1}$.

AZTTP binding and incorporation were also globally fit to a two-step nucleotide binding model. The ground state binding of AZTTP is much tighter that than of TTP, with a dissociation constant of 26±2 μM (Figure 2.8A/B). Global analysis resulted in a value of $k_2$ = 105±11 $s^{-1}$, governing conformational change from the "open" to "closed" state. Since this rate is low enough to be accurately determined at 37°C, Arrhenius analysis was no longer required. The reverse rate of the conformational change ($k_{-2}$) was fit to a value of 48.4±3.7 $s^{-1}$, as measured by the competition assay (Figure 2.8C). And it is worth noting that the amplitude dependence of the stopped flow binding experiment (Figure 2.8A) provided additional constraint on $k_{-2}$ since equilibrium was established prior to chemistry due to the fact that $k_{-2}$ was 4-fold larger than $k_3$. The rate of chemistry was 11.0±0.6 $s^{-1}$ (Figure 2.8A/B/D). PPi release assay afforded a rate of 1000±650 $s^{-1}$, with a lower limit of 172 $s^{-1}$ (Figure 2.8D). Global fitting results for TTP and AZTTP binding and incorporation are summarized in Table 2.3.

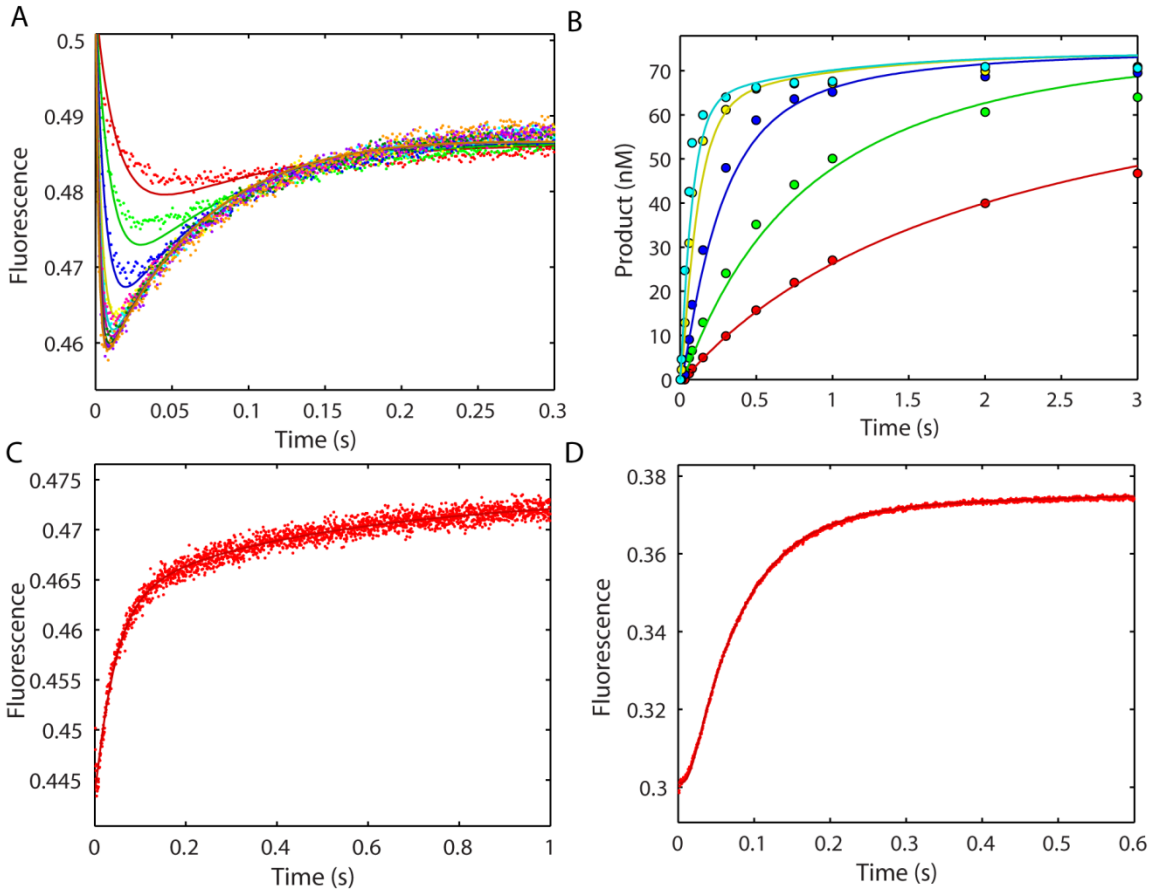



**Figure 2.5** *Global fitting of AZTTP Binding and Incorporation by MDCC-labeled HIV RT_T69SSS.* (A) AZTTP concentration dependence of nucleotide binding and incorporation. Various concentrations of AZTTP (2.5, 5, 10, 20, 30, 40, 60, 80, and 100 μM) were mixed with a pre-formed enzyme-DNA complex (150 nM 25/45mer and 100 nM HIVRT_T69SSS-MDCC). (B) Rapid quench flow AZTTP concentration dependence of the rate of product formation. Various concentrations of AZTTP (0.1, 0.25, 0.75, 3, 15 μM) were mixed with a pre-formed enzyme-DNA complex (75 nM 25/45mer DNA and 175 nM HIVRT_T69SSS-MDCC). (C) The rate of re-opening of the 'closed' ternary complex was measured by rapidly mixing a pre-formed enzyme-$DNA_{dd}$-AZTTP complex (200 nM HIVRT_T69SSS-MDCC, 300 nM 25ddA/45mer, 2.5 μM nucleotide) with a nucleotide trap consisting of 2.5 μM unlabeled enzyme-DNA complex. (D) Rate of PPi release during the incorporation of AZTTP was monitored by rapidly mixing 25 μM of AZTTP with a pre-formed enzyme-DNA complex (100nM 25/45mer DNA and 150 nM HIV RT_T69SSS-MDCC) in the presence of 1.5 μM of MDCC-PBP and 0.6 μM of PPase. In each panel the smooth lines show the best fit achieved by globally fitting panels A-D simultaneously. Rate constants are summarized in Table 2.3.

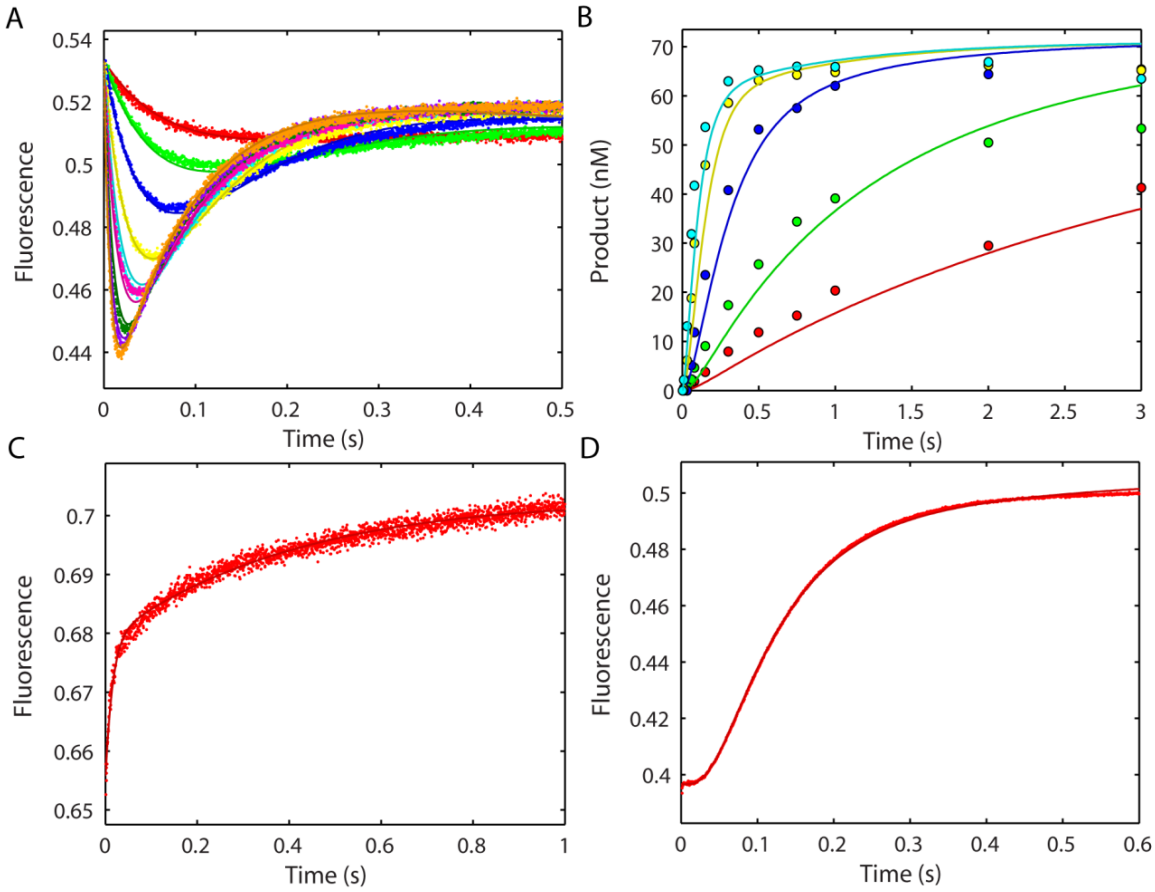



**Figure 2.6** *Global fitting of TTP Binding and Incorporation by MDCC-labeled HIV RT_TAMs_T69SSSVD.* (A) TTP concentration dependence of nucleotide binding and incorporation. Various concentrations of TTP (2.5, 5, 10, 20, 30, 40, 60, 80, and 100 μM) were mixed with a pre-formed enzyme-DNA complex (150 nM 25/45mer and 100 nM HIVRT_TAMs_T69SSSVD-MDCC). (B) Rapid quench flow TTP concentration dependence of the rate of product formation. Various concentrations of TTP (0.25, 0.75, 3, 10, 30 μM) were mixed with a pre-formed enzyme-DNA complex (75 nM 25/45mer DNA and 175 nM HIVRT_TAMs_T69SSSVD-MDCC). (C) The rate of re-opening of the 'closed' ternary complex was measured by rapidly mixing a pre-formed enzyme-$DNA_{dd}$-TTP complex (200 nM HIVRT_TAMs_T69SSSVD-MDCC, 300 nM 25ddA/45mer, 9 μM TTP) with a nucleotide trap consisting of 9 μM unlabeled enzyme-DNA complex. (D) Rate of PPi release during the incorporation of TTP was monitored by rapidly mixing 25 μM of TTP with a pre-formed enzyme-DNA complex (100nM 25/45mer DNA and 150 nM HIVRT_TAMs_T69SSSVD-MDCC) in the presence of 1.5 μM of MDCC-PBP and 0.6 μM of PPase. In each panel the smooth lines show the best fit achieved by globally fitting panels A-D simultaneously. Rate constants are summarized in Table 2.3.

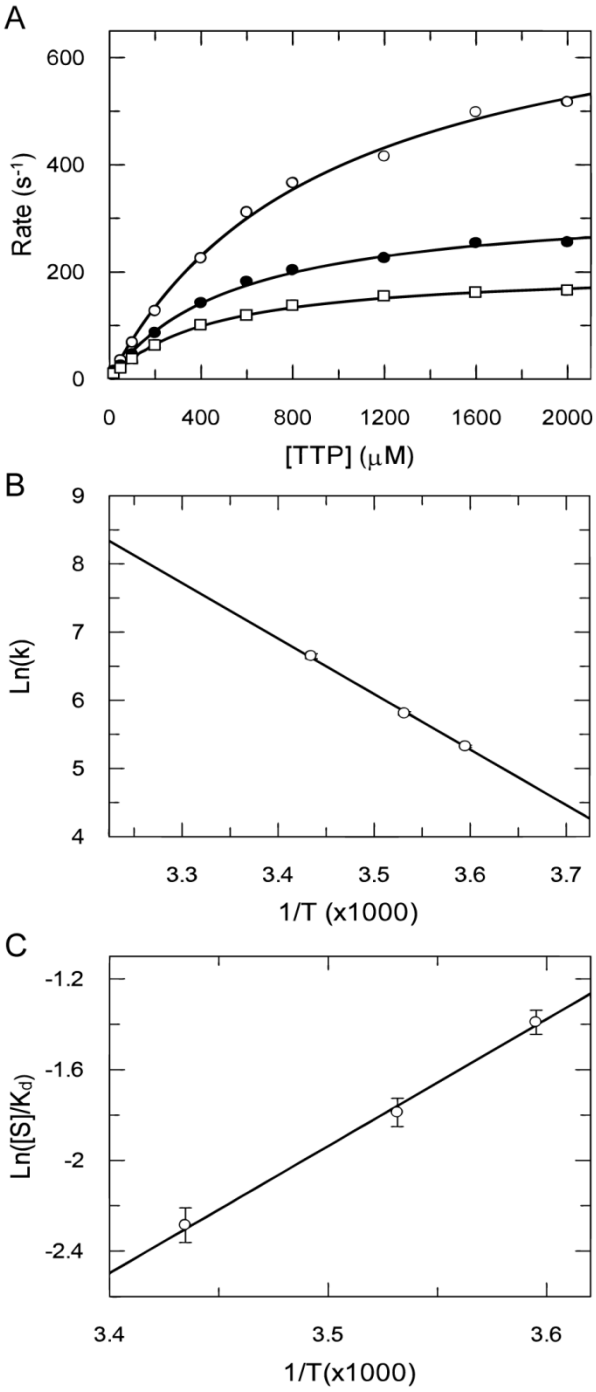



**Figure 2.7** *Temperature Dependence of the HIVRT TAMs_T69SSSVD Conformational Change Rate upon TTP Binding.* (A) The experiment shown in Figure 2.6A was repeated at various temperatures. A pre-incubated complex of 150 nM 25/45mer and 100 nM MDCC-labeled HIVRT TAMs_ T69SSSVD was rapidly mixed with various concentrations of TTP (20, 50, 100, 200, 400, 600, 800, 1200, 1600 and 2000 μM) at 5°C (□), 10°C (●), and 18°C (○). The concentration dependence of the fast decrease of fluorescence (corresponding to conformational closing) was fit to hyperbolic equation to obtain the maximum rate of the conformational change ($k_2$) at each temperature resulting in rates of 205±3, 332±9 and 770±29 $s^{-1}$; as well as estimates of $K_d$ = 402±2, 600±2, and 980±2 μM at 5, 10, and 18°C, respectively. (B) The temperature dependence of $k_2$ was analyzed by Arrhenius plot to estimate a maximum rate for $k_2$ of 4350±380 $s^{-1}$ at 37°C. The value was then fixed to constrain data fitting by global analysis of the data in Figure 2.6. (C) The temperature dependence of $K_d$ was analyzed by Van't Hoff plot to derive estimates of ΔH of -45.9±2.6 kJ/mol, ΔS of 0.18±0.01 kJ $K^{-1}$/mol, as well as ΔG of 3.2±0.1, 4.2±0.1, and 5.5±0.1 kJ /mol at 5, 10, and 18°C, respectively.

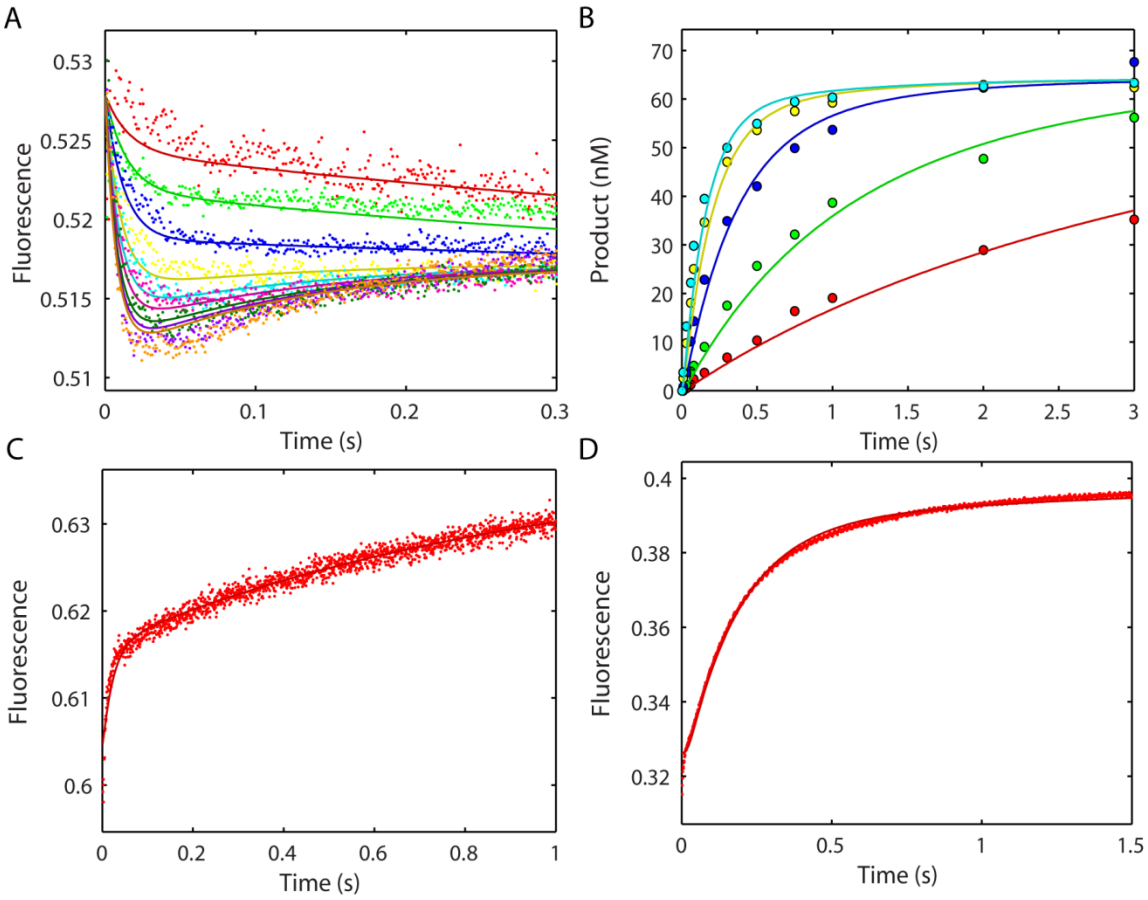



**Figure 2.8** *Global fitting of AZTTP Binding and Incorporation by MDCC-labeled HIV RT_TAMs_T69SSSVD*. (A) AZTTP concentration dependence of nucleotide binding and incorporation. Various concentrations of TTP (2.5, 5, 10, 20, 30, 40, 60, 80, and 100 μM) were mixed with a pre-formed enzyme-DNA complex (300 nM 25/45mer and 200 nM HIVRT_TAMs_T69SSSVD-MDCC). (B) Rapid quench flow AZTTP concentration dependence of the rate of product formation. Various concentrations of AZTTP (0.5, 1.5, 6, 20, 60 μM) were mixed with a pre-formed enzyme-DNA complex (75 nM 25/45mer DNA and 175 nM HIVRT_TAMs_T69SSSVD-MDCC). (C) The rate of re-opening of the 'closed' ternary complex was measured by rapidly mixing a pre-formed enzyme-$DNA_{dd}$-AZTTP complex (200 nM HIVRT_TAMs_T69SSSVD-MDCC, 300 nM 25ddA/45mer, 10 μM AZTTP) with a nucleotide trap consisting of 10 μM unlabeled enzyme-DNA complex. (D) Rate of PPi release during the incorporation of AZTTP was monitored by rapidly mixing 25 μM of AZTTP with a pre-formed enzyme-DNA complex (100nM 25/45mer DNA and 150 nM HIVRT_TAMs_T69SSSVD-MDCC) in the presence of 1.5 μM of MDCC-PBP and 0.6 μM of PPase. In each panel the smooth lines show the best fit achieved by globally fitting panels A-D simultaneously. Rate constants are summarized in Table 2.3.

**Determination of Specificity Constants Governing TTP and AZTTP Binding and Incorporation to T69SSS and TAMs_T69SSSVD HIVRTs**

The specificity constant ($k_{cat}/K_m$) was calculated as $k_{cat}/K_m = k_1k_2k_3/ (k_2k_3+k_{-1}(k_{-2}+k_3))$ for each combination of nucleotide and enzyme. The specificity constants defining TTP and AZTTP incorporation by the HIVRT_T69SSS mutant were 6.8±0.6 and 7.3±0.7 $\mu M^{-1}s^{-1}$, respectively. Based on these values a net discrimination ($D = [k_{cat}/K_m]_{TTP}/[k_{cat}/K_m]_{AZTTP}$) of 0.93±0.11 was derived (Table 2.4). The net $K_d$ values for the two-step nucleotides binding ($K_{d,net} = 1/(K_1(1+K_2))$) were 0.20±0.02 and 1.3±0.1 μM, respectively.

The specificity constants defining TTP and AZTTP incorporation by the HIVRT_TAMs_T69SSSVD mutant were 1.4±0.1 and 0.7±0.1 $\mu M^{-1}s^{-1}$, respectively. These values defined the AZTTP discrimination by the TAMs_T69SSSVD mutant of 2.0±0.3, a modest increase from the T69SSS mutant. Except for reduced specificity constants for TTP and AZTTP, introduction of fingers insertion into the TAMs background also resulted in weaker net nucleotide binding affinities of 1.0±0.1 and 11.8±1.9 μM, respectively. The reduction in the overall nucleotide binding interaction for the TAMs_T69SSSVD mutant was 5~9 fold. This mutant also showed a reduction of 5-fold in enzyme efficiency, as defined by $k_{cat}/K_m$ values.

**Effect of Fingers Insertion on AZTMP Excision and the Binding of the Next Correct Nucleotide**

ATP dependent excision of AZT-monophosphate (AZTMP) from the 3' end of the blocked primer has been identified as a significant source of resistance to AZT (*122*).

In this study, ATP concentration dependent excision and extension assays were conducted to determine efficiencies ($k_{exc}/K_{m,ATP}$) of the RTs (WT/T69SSS/TAMs/TAMs_T69SSSVD) to rescue AZTMP-terminated primer (Figure 2.9A-D). Since AZTMP-terminated primer rescue can be specifically inhibited by the next complementary nucleotide (*126-127*), dCTP (the next correct nucleotide) concentration dependence of excision rates by various RTs were determined to explore if the fingers insertion have any effect on the inhibitory effect of dCTP (Figure 2.10A-D). Global analyses were performed independently for WT (Figure 2.9A/2.10A), T69SSS (Figure 2.9B/2.10B), TAMs (Figure 2.9C/2.10C), and TAMs_T69SSSVD (Figure 2.9D/2.10D), according to Scheme 2.4, as illustrated by the smooth curves in each figure, in order to derive the rate constants that were used to calculate the excision efficiencies ($k_{exc}/K_{m,ATP}$) and the apparent binding affinities of dCTP ($K_{d,dCTP}$), as summarized in Table 2.5.

**Table 2.5** *Excision Efficiencies of AZTMP and Apparent Dissociation Constants for the Binding of the Next Correct Nucleotide*

| Enzyme | $k_{exc}/K_{m,ATP}$ ($M^{-1}s^{-1}$) | Fold Effect | $K_{d,dCTP}$ ($\mu M$)[c] | Fold Effect |
|---|---|---|---|---|
| WT | 0.40±0.13 [a] | | 273±35 | |
| T69SSS | 0.67±0.17 [a] | 1.7X | 370±35 | 1.3X |
| TAMs | 1.95±0.05 [b] | 4.9X | 720±90 | 2.6X |
| TAMs_T69SSSVD | 0.74±0.16 [b] | 1.8X | 1380±170 | 5.0X |

[a] $k_{exc}/K_{m,ATP}$ values for WT and T69SSS RTs were calculated based on $k_{exc}$ and $K_{m,ATP}$ values derived from global analysis according to scheme 2.4 using $K_{m,ATP}$ values obtained from the conventional fitting as starting estimates.

[b] $k_{exc}/K_{m,ATP}$ values for those two mutants were derived from global analysis of the data fitted to Scheme 2.4 allowing for a one-step excision.

[c] The apparent $K_d$ as a result of a reverse translocation step prior to dCTP binding was calculated as $K_{d,1}(1+K_0)$, and in each case $K_0$ equals 6.

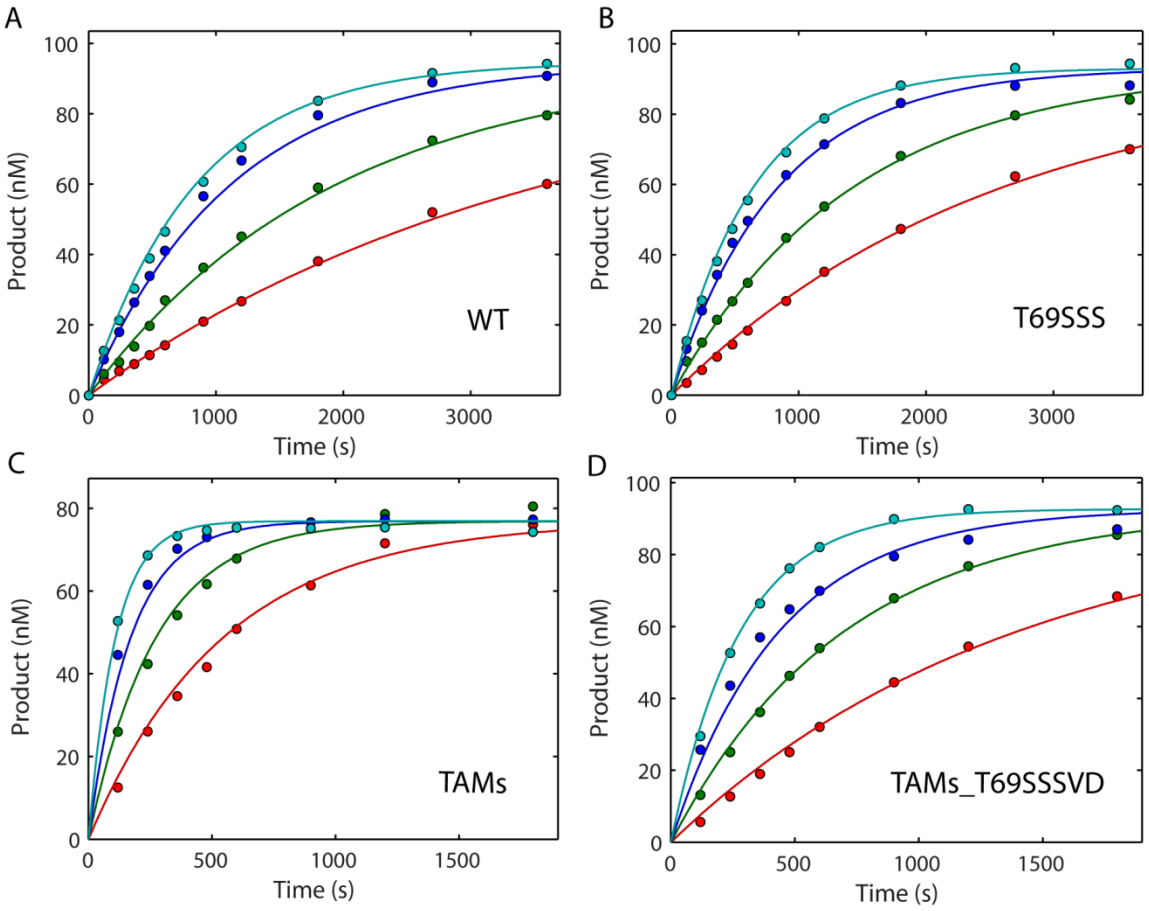



**Figure 2.9** *ATP concentration dependent excision of AZTMP by various RTs.* ATP concentration dependence of AZTMP excision was measured by manual chemical quench flow method. A pre-formed enzyme-$DNA_{AZTMP}$ complex (150 nM WT RT (A) or T69SSS (B), 100nM 25/45mer, 12.5 μM AZTTP) was rapidly mixed with 100 μM TTP and 1 μM dCTP in the presence of various concentrations of ATP (0.8, 1.6, 3.2, and 4.8 mM). Similarly, a pre-formed enzyme-$DNA_{AZTMP}$ complex (150 nM TAMs (C) or TAMs_T69SSSVD (D), 100 nM 25/45mer, 12.5 μM AZTTP) was rapidly mixed with 100 μM TTP and 1 μM dCTP in the presence of various concentrations of ATP (1.0, 2.0, 3.2, and 5.0 mM).

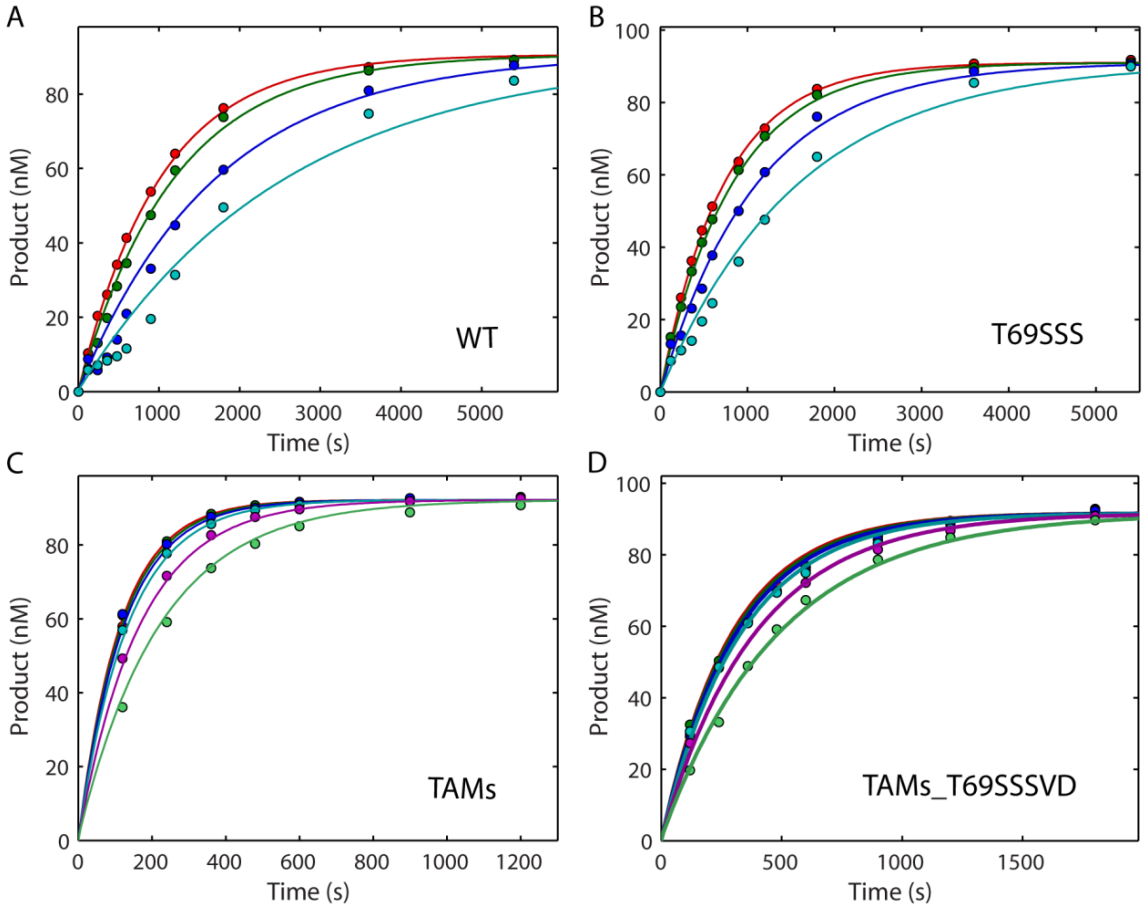



**Figure 2.10** *dCTP concentration dependent excision of AZTMP by various RTs.* dCTP concentration dependence of AZTMP excision was measured by manual chemical quench flow method. A pre-formed enzyme-$DNA_{AZTMP}$ complex (150 nM WT (A), or T69SSS (B), 100 nM 25/45mer, 12.5 μM AZTTP) was rapidly mixed with 100 μM TTP and 4.8 mM ATP in the presence of various concentrations of dCTP (100, 200, 500, and 1000 μM). Similarly, a pre-formed enzyme-$DNA_{AZTMP}$ complex (150nM TAMs (C), or TAMs_T69SSSVD (D), 100 nM 25/45mer, 12.5 μM AZTTP) was rapidly mixed with 100 μM TTP and 4.8 mM ATP in the presence of various concentrations of dCTP (10, 50, 100, 200, 500, and 1000 μM). Global analysis were performed independently for WT (Figure 2.9A/2.10A), T69SSS (Figure 2.9B/2.10B), TAMs (Figure 2.9C/2.10C), and TAMs_T69SSSVD (Figure 2.9D/2.10D), according to Scheme 2.4. Excision efficiencies ($k_{exc}/K_{m,ATP}$) and apparent binding affinities of dCTP ($K_{d,dCTP}$) resulting from global analysis of kinetic data are summarized in Table 2.5.

Conventional data analysis of ATP concentration dependence of excision rates of WT RT afforded a $K_{m,ATP}$ value of 5250±2130 μM. In addition, our unpublished data shown that ~86% of the AZTMP blocked primers reside at the N site, from which an equilibrium constant ($K_0$) of 6 was obtained. These values were set as the starting estimates in the global fitting to derive other rate constants. The rate of excision ($k_{exc}$) was 0.0024±0.0008 $s^{-1}$, based on which a $k_{exc}/K_{m,ATP}$ of 0.40±0.13 $M^{-1}s^{-1}$ were derived (Figure 2.9A). Global analysis resulted in a value of 39±5 μM ($K_{d,1}$) governing the ground-state binding of dCTP to RT locating at the P site. Equilibrium constant of the reverse translocation between the P site and N site ($K_0$) was fixed at a value of 6, as mentioned above. These values produced an apparent $K_d$ ($K_{d,1}(1+K_0)$) of dCTP binding as 273±35 μM (Figure 2.10A). Similar experiments and data analysis were performed to determine kinetics governing AZTMP excision by and dCTP binding to the T69SSS mutant. Estimate of $K_m$ for ATP as 4000±2160 μM was derived from the conventional fitting, and were then locked in the global data analysis, resulting in a $k_{exc}$ value of 0.0028±0.0007 $s^{-1}$, which produced a $k_{exc}/K_{m,ATP}$ of 0.67±0.17 $M^{-1}s^{-1}$ (Figure 2.9B). Data contained within the families of curves for dCTP and ATP concentration dependence of excision rates were used to constrain a dCTP binding constant ($K_{d,1}$) of 53±5 μM. The binding of dCTP to T69SSS was 1.4-fold weaker than to WT RT, with apparent $K_d$ of 370±35 μM (Figure 2.10B).

In cases of TAMs and TAMs_T69SSSVD, estimates of $K_m$ values for ATP were not available since binding of ATP to these mutants was very weak and therefore the

maximum concentration of 5 mM ATP was too low to even partially saturate the reaction. Consequently, a one-step excision mechanism was used to derive the $k_{exc}/K_{m,ATP}$ ($k_{eff}$) values directly, which afforded 1.95±0.05 and 0.74±0.16 $M^{-1}s^{-1}$ for the TAMs and TAMs_T69SSSVD mutants, respectively (Figure 2.9C/D). Similarly, equilibrium constant governing the reverse translocation between the P site and N site ($K_0$) was fixed at 6. The apparent binding affinity ($K_{d,dCTP}$) of dCTP to the TAMs mutant was 720±90 μM, as calculated from the ground-state $K_d$ ($K_{d,1}$) of 103±13 μM and the equilibrium constant ($K_0$) of 6 (Figure 2.10C). Binding of dCTP to TAMs_T69SSSVD mutant was 1.9-fold weaker than to TAMs RT, with an apparent $K_d$ of 1380±170 μM, derived from a $K_{d,1}$ of 197±24 μM and a $K_0$ of 6 (Figure 2.10D).

**MDCC Labeling Does Not Affect the Excision Activity of HIV RT**

Because mutants with fingers insertion were MDCC labeled control experiments were performed to test if this labeling has any effect on the excision reaction (Figure 2.11). WT and WT-MDCC were tested at 3.2 mM of ATP, yielding comparable excision rates as 0.0012±0.00005 and 0.0013±0.00004 $s^{-1}$, respectively, implying that the labeling did not affect enzyme activity. Previous data also showed identical kinetics of incorporation by WT HIVRT with or without MDCC labeling (*86*).

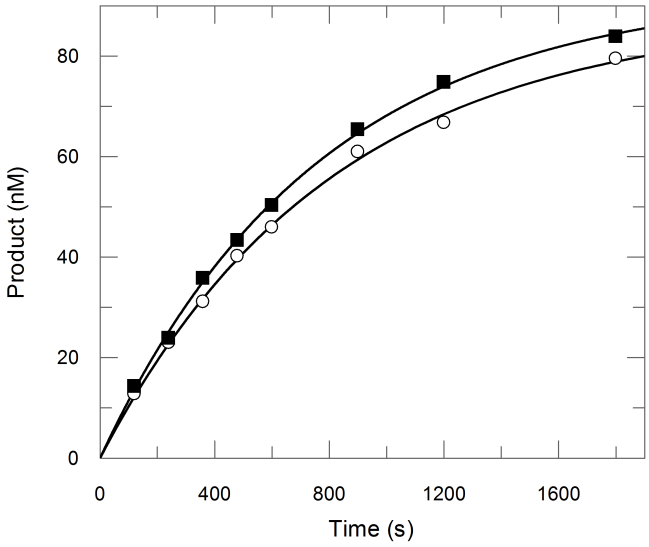



**Figure 2.11** *Excision Activities of WT and WT-MDCC HIVRT Proteins*. The excision activities by WT (open circle) and WT-MDCC (solid square) HIVRT are compared at 3.2 mM of ATP. A pre-formed enzyme-$DNA_{AZTMP}$ complex (150 nM WT RT or WT-MDCC, 100nM 25/45mer, 12.5 μM AZTTP) was rapidly mixed with 100 μM TTP and 1 μM ddCTP in the presence of 3.2 mM ATP. Excision data were fit to a single exponential equation. Excision rates obtained with WT and WT-MDCC HIVRT proteins are 0.0012±0.00005 and 0.0013±0.00004 $s^{-1}$, respectively.

## 2.4 DISCUSSION

The effects of fingers insertion of HIVRT on nucleotide specificity and ATP mediated excision of AZT have been examined. Previous studies suggested that fingers insertions might have only limited effect on discrimination against NRTIs. However, structural modeling showed that insertions reside in the vicinity of and therefore can affect the dNTP-binding pocket (*109*). Meanwhile, single dNTP exclusion assays confirmed that insertions in combination of AZT resistant mutations could reduce misincorporation efficiencies to a large extent (*128*). Our data also showed that a Ser-Ser insertion alone can reduce misincorporation efficiencies by 10~15 fold, indicating that insertions might have some influence on nucleotide binding and/or incorporation (An Li, unpublished data). Given those considerations, the problem was revisited in this research by pre-steady-state experiments and global analysis of the data to a comprehensive model

shown in Scheme 2.1, 2.2 and 2.3. In addition we have examined the role of the conformational change in nucleotide discrimination. It has already been shown that fitting data based upon numerical integration of the rate equations with no simplifying assumptions is far more robust and demanding than fitting to equations derived from a simplified form of the model because it requires that the model account for both the rate and amplitude of all reactions simultaneously (*77, 86*). In the previous studies, no information existed to define the rate of pyrophosphate release ($k_4$). Although available evidence indicates that pyrophosphate (PPi) release is usually fast, it was not known if fingers insertion would affect PPi release. The first three experiments (e.g., Figure 2.2A-C) did not provide enough information to distinguish between two alternative mechanisms: either slow PPi release (18 $s^{-1}$) with faster chemistry (33 $s^{-1}$, data not shown) or fast PPi release (>200 $s^{-1}$) with slower chemistry (21 $s^{-1}$). To address this question, the rate of PPi release was directly measured using a coupled fluorescence assay previously developed in our lab, which at last supported the fast PPi release model, indicating that fingers insertion mutations did not affect PPi release. Notably, a more accurate estimation of rate of chemistry could be obtained since additional information from this assay could provide more constrains in fitting the data globally.

Discrimination factors for WT RT (1.02±0.08) and TAMs mutant (2.4±0.05) were obtained in our previously studies (*77*). The comparisons (WT *vs.* T69SSS and TAMs *vs.* TAMs_T69SSSVD) indicate that a Ser-Ser insertion did not improve discrimination against AZT either in an otherwise WT RT or in the TAMs mutant. As reported previously, the small 2-fold improvement in discrimination may not be enough to provide

the clinically observed high level of AZT resistance conferred by the 69 insertion complex. Although certain published data have shown similar trends (*119, 124*), the steady-state kinetic analysis failed to provide the mechanistic basis for this absence of discrimination. Results shown in the current study, however, provided information to address this question. For the T69SSS insertion mutants, the ground-state binding of TTP ($1/K_1$, Table 2.3) was much weaker (10-100 fold) than that for AZT. The same pattern was also observed in our previous studies conducted with WT and TAMs proteins, but only by a factor of 3, indicating that the Ser-Ser insertion improved the ground state binding of AZTTP to a much larger extent than the TAMs mutant. Although ground-state binding is reduced, TTP still binds 6~12 folds tighter than AZTTP (Table 2.4, $K_{d,net}$) as a result of contributions from the much more favorable conformational change step. It's worth noting that the insertion greatly reduced the equilibrium constants ($K_2$) of the isomerization induced by AZTTP. In the case of TAMs_T69SSSVD, a $K_2$ value around 2.2 was obtained. This small equilibrium constant might account for the small amplitudes of fluorescent signals observed in Figure 2.8A. Arrhenius analysis to determine the maximum rate of the conformational change ($k_2$) at 37°C was performed only for the binding of TTP because rates of AZTTP resided in the detectable range of the instrument. The negative $\Delta H$ values between 45.9 and 48.4 kJ/mol indicate initial "ground-state" binding of TTP is an exothermic process (Figure 2.3 and 2.7).

We have shown the role of induced-fit in determining enzyme specificity to nucleotides in the cases of T7 DNA polymerase (*87*) as well as HIV RT (*77, 86*). Briefly, the relative rates of chemistry ($k_3$) and the reverse of the conformational change ($k_{-2}$)

determines if the equilibrium is established prior to chemistry or not. Based on different kinetic partitioning, different simplified derivations of $k_{cat}/K_m$ ($K_1k_2$ or $K_1K_2k_3$) could be obtained. However, the data in this paper did not follow this simple trend because observed differences in these rates ($k_{-2}$ vs $k_3$) are not significant enough to make the simplifications. We observe that $k_{cat}/K_m$ values for T69SSS incorporation of TTP and AZTTP are equal in spite of the weaker ground-state binding of TTP ($K_{d,TTP} >> K_{d,AZTTP}$), because it is offset by a more favorable conformational change for TTP ($K_{2,TTP} >> K_{2,AZTTP}$), while the rates of chemistry are comparable. A similar analysis of TAMs_T69SSSVD provides the mechanistic basis for the modest increase in discrimination in spite of weaker ground-state binding of TTP to the mutant enzyme than AZTTP. Although TTP binds much more weakly to the mutant enzyme (>100 folds weaker), the conformational change to the "closed" state is more highly favored compared to AZTTP induced isomerization ($K_{2,TTP}$ = 2900±320, $K_{2,AZTTP}$ = 2.2±0.3), while the rates of chemistry are equal. Consequently, the whole process is similar to that of T69SSS but the changes in ground-state $K_d$ and equilibrium of the conformational change are more profound. This can also be seen in the free energy profiles for binding and incorporation of nucleotides (Figure 2.12). It is notable that in the cases of AZTTP incorporation by either T69SSS or TAMs_T69SSSVD, the reverse conformational change rates are higher than rates of chemistry ($k_{-2} > k_3$). This allows the two-step binding mechanism to come to equilibrium, leading to reduced $K_m$ values, which reflect the true $K_{d,net}$ (Table 2.4). Consequently, these reduced $K_m$ values contribute to the absence of significant discrimination against AZTTP. This absence of discrimination

against AZT is likely due to the high degree of structural homology between the inhibitor and TTP. It was reported that human mitochondrial DNA polymerase discriminates against AZT by slowing the rate of PPi release after incorporation (*90*) while our data clearly showed that fingers insertion did not affect PPi release.

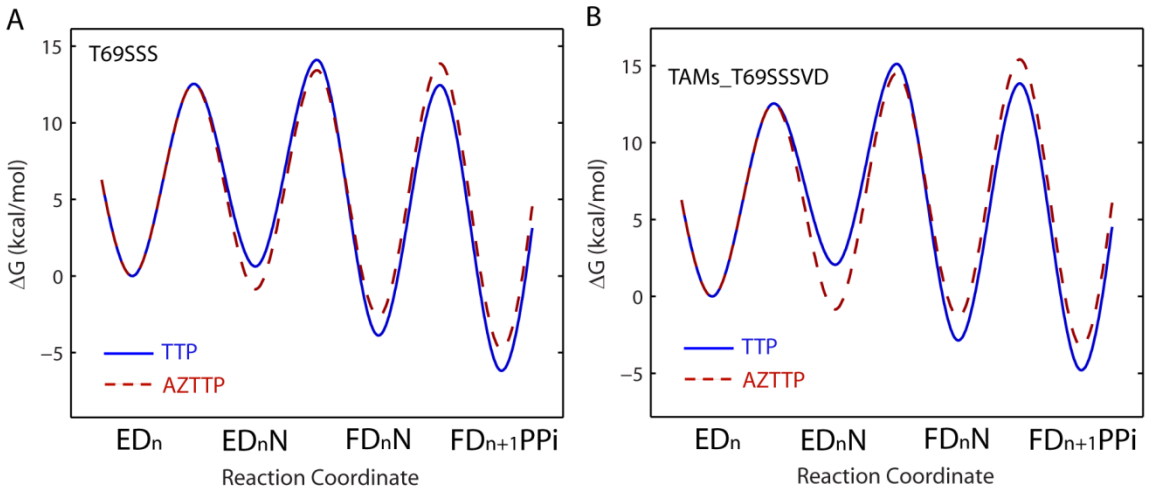



**Figure 2.12** *Free-energy profiles for HIV-1 RT mutants.* (A) Free-energy diagram for TTP and AZT nucleotide binding and incorporation reactions by T69SSS mutant. The free energy was calculated as $\Delta G = RT[\ln(kT/h)-\ln(k_{obs})]$ kcal/mol using rate constants from Table 2.3. The constant $k$ is the Boltzmann constant, $T$ is 310K, $h$ is Planck's constant and $k_{obs}$ is the first-order rate constant. The nucleotide concentration was set as 100 μM. (B) Free-energy diagram for TTP and AZT nucleotide incorporation reactions by the TAMs_T69SSSVD mutant.

Several findings have demonstrated that drug resistance associated with the 69 insertion complex also involves ATP-mediated excision reactions and suggested that the fingers insertion enhances this reaction by reducing inhibition from the next correct nucleotide (*116-117, 119-120, 124*). ATP concentration dependent primer rescue and extension assays were performed to monitor effects of fingers insertion on ATP binding

during excision reaction. Meanwhile, the 50% inhibitory concentration ($IC_{50}$) analysis widely used in many studies is not an accurate measurement of the binding affinity of the next cognate nucleotide due to the complexity of the system. Therefore, dCTP (the next complementary nucleotide in this experimental set up) concentration dependent excision assays were also performed and globally fit to a comprehensive model shown in Scheme 2.4. There is a crucial translocation step (between N site and P site) in this model, the equilibrium of which determines fates of the blocked primer: either dead end complex (DEC) formation or rescue and extension. Our single molecule analysis of pulse width lifetime (nucleotide bound state) and interpulse duration (nucleotide unbound state) demonstrated that only 14% of AZTMP terminated primers reside at P site that are capable of nucleotide binding, from which an equilibrium constant ($K_0$) of 6 was obtained for WT RT (unpublished data). This number was used as a first estimation of $K_0$ in global fitting, which was then fixed for better constrain on other parameters. Unfortunately, we did not have any information regarding $K_0$ for RT mutants. Site-specific hydroxyl foot-printing was thought to be an ideal method to assess subtle structural differences at single nucleotide resolution (*72, 129-130*). In order to explore the equilibrium of the translocation of AZTMP terminated primers, foot-printing assays were conducted using WT RT, TAMs and TAMs_T69SSSVD mutants. As expected, all AZTMP-terminated primers predominantly resided at the N site no matter WT or mutant proteins were used (data not shown). However, this method is hardly an ideal one for accurate determination of the translocational equilibrium for two reasons: 1) the reaction is too slow compared to rapid equilibration, and 2) different reducing reagents (ascorbate

or DTT) yield various results, which suggests this method is not accurate. In studies of fluorescence changes induced by the binding of dCTP to ddTMP or d4TMP terminated primers, we found that the ratio of the fraction of fast conformational change to that of the slower isomerization provided sufficient information to define the equilibrium constant. The results suggested that the TAMs mutations shifted the equilibrium to favor the residence of DNA at the N site by 5-20 fold (Chapter 3). However, no signal was observed in similar experiments conducted with AZTMP-terminated primers, probably because both azido group and TAMs mutations facilitate the primers to reside at N site where the incoming nucleotide cannot bind. Based on these observations, it is known $K_0$ has a lower limit of 6 which was used in global fitting for all RTs used in this study. On the other hand, the apparent dissociation constants for dCTP ($K_{d,dCTP}$) were determined by the product of $K_{d,1}(1+K_0)$, which was well defined in spite of uncertainties in $K_{d,1}$ and $K_0$. An estimation of $K_0$ as 6 is acceptable which can then be fixed in the global analysis so that well constrained parameters governing dCTP binding and AZTMP excision were obtained (Table 2.5).

Fingers insertion promoted excision efficiency modestly when introduced into WT RT. To our surprise, the presence of Ser-Ser insertion in combination with TAMs mutations decreased the excision efficiency by a factor of ~2.5. Others proposed that fingers insertions enhanced ATP-dependent excision because it reduced inhibition from the next cognate nucleotide. However, $K_{d,dCTP}$ derived for TAMs and TAMs_T69SSSVD are 720±90 and 1380±170 μM, respectively. This weak binding shouldn't make any significant difference under physiological concentrations of nucleotides which are in the

5-40 μM range (*131*). Also the previously reported $IC_{50}$ values for WT (9 μM) and T69SSS (12 μM) RTs (*119*) are underestimated since most of AZTMP terminated primers locate at the N site (no matter WT or mutant proteins were used) where the next complementary nucleotide cannot bind, which would make the apparent binding much weaker. Other groups reported higher $IC_{50}$ values of 80±13 (*132*), 110±40 (*122*) or > 500 μM (*78*) for WT protein, which are more close to our $K_d$ values. Also, fixed time point analysis is not convincing to draw the conclusion that fingers insertion can enhance ATP-mediate excision. Others did observe an increase of excision rate but failed to consider effects on ATP binding affinity. Although we do not have enough information to define ATP binding affinities to AZTMP terminated primer for the two mutants containing TAMs mutations since the highest concentration we could reach is 5 mM, above which excision rate would decrease. Nonetheless, our data refute claims of increased ATP binding affinity for the TAMs mutants. One possible scenario is that the Ser-Ser insertion increases maximum excision rate but makes ATP binding weaker and the combination of these two effects gives a smaller $k_{exc}/K_{m,ATP}$ value. Actually, fingers insertion weakened the binding of ATP by 1.4-2.4 fold in experiments studying the excision of ddTMP and d4TMP (Chapter 3).

Notably, the TAMs mutations enhanced the excision efficiency only by a factor of 5, which is smaller than some previously published values. However, excision efficiencies of our mutants (e.g., $2\,M^{-1}\,s^{-1}$ for TAMs) are comparable to previously reported ones, but the $k_{eff}$ of our WT RT is much higher than those in the literature (e.g., only ~10% completion within 30 min (*116*)). One possible explanation is that the

untagged RT proteins we made are much more efficient than those his-tagged ones, and therefore harboring faster polymerization (7 vs. 0.5 $\mu M^{-1}$ $s^{-1}$ (*119*)), as well as excision. Also, some other groups reported 5-fold increase of excision efficiency by the TAMs mutant, for example, 0.33 *vs.* 0.066 $M^{-1}s^{-1}$ (*132*) and 0.39 *vs.* 0.08 $M^{-1}s^{-1}$ (*122*).

In conclusion, our data suggested that the 69 insertion complex exhibited only moderate discrimination against AZT incorporation, as well as modest increase in ATP-mediate excision, a finding that is not consistent with a widely accepted proposal that ATP-dependent phosphorolytic activity plays a predominant role in AZT resistance. The fingers insertions may increase the maximum excision rate but weaken ATP binding to a larger extent. Although binding of the next cognate nucleotide to an AZTMP terminated primer is also weaker for the 69 insertion complex, the binding is too weak to make any significant contribution to excision activities at physiological conditions. Consequently, the two mechanisms seem to be equally important and work together during DNA-dependent DNA synthesis or probably, during the RNA-dependent DNA synthesis also. Given the size of the viral genome, the two mechanisms can take place at multiple sites, which might explain why the relatively modest changes can result in high level of drug resistance. Further studies of effects of fingers insertions during other stages of reverse transcription, e.g., strand transfer (*133*) and strand displacement synthesis, or their influence in viral fitness may provide a more detailed basis for evolving advantages of the 69 insertion complex.

# Chapter 3: Effects of AZT Resistance Mutations and Dipeptide Insertions in Fingers of HIV-1 Reverse Transcriptase on ATP-Dependent Excision of Thymidine Analogs

## 3.1 INTRODUCTION

HIV-1 reverse transcriptase (HIVRT) replicates the viral RNA genome through its RNA and DNA-dependent DNA polymerization and RNase H degradation activities at the early stage of the viral cycle (*134-135*). Consequently, HIVRT has been a major target for the development of antiretroviral therapies (*60-61*). Nucleoside/nucleotide RT inhibitors (NRTIs) and nonnucleoside RT inhibitors (NNRTIs) are two distinct classes of compounds targeting HIVRT, the former are a group of structurally diverse analogs of the natural substrates for DNA synthesis, which lack the 3'-OH and thereby act as chain terminators after incorporation by RT. Among these compounds, AZT (zidovudine) and d4T (stavudine) are thymidine analogs, the former was the first approved by the US Food and Drug Administration (FDA) in 1987 and has been widely used until now while d4T is no longer used in therapies due to toxic side effects (*63-64, 136*).

Drug resistance is a major challenge in HIV therapies since sustained viral suppression requires prolonged usage of medications (*66*). Resistance mutations for all approved NRTIs have been identified and certain combinations of mutations were found to confer cross-resistance to all (or nearly all) NRTIs (*67*). TAMs mutations (M41L, D67N, K70R, L210W, T215F/Y, and K219Q/E) confer high levels of resistance to AZT, as well as lower levels of resistance to all other NRTIs (*67, 108, 137-138*), and therefore are an example of multi-NRTI resistance. Several studies were performed to explore the molecular mechanism of AZT resistance conferred by TAMs mutations and the most

accurate evidence indicates that TAMs mutations confer only a twofold change in specificity relative to incorporation of TTP (*77, 139-140*). Rather, these mutations provide an enhanced ATP-mediated removal of chain terminators from the 3' end of the blocked primer (*79, 122, 126*). However, our previous studies reported that the TAMs mutations enhanced the excision efficiency only by a factor of 5 (Chapter 2), but also reduce the binding of the next correct nucleotide which would otherwise block the ATP-mediated excision, and a recent single turnover kinetic analysis showed that discrimination against AZTTP incorporation was increased by 2.4-fold (*77*). Therefore, contrary to prior reports, the clinically observed 200-fold resistance to AZT appears to result from the combined exclusion and excision mechanisms operating over multiple sites during replication.

Boyer *et al.* proposed a model of ATP-mediated NRTI excision and its inhibition by the next correct nucleotide based on both biochemical and structural evidences (*79*). Central to this model is the finding that the 3' end of the primer is in dynamic equilibrium between the nucleotide-binding site (N site) and the priming site (P site). For the excision reaction to be carried out, the chain terminator must reside in the N site where nucleophilic attack by ATP can occur. When the 3' end of a primer resides in the P site, the N site is open for the binding of the next cognate nucleotide. In this situation, a "dead-end complex" (DEC) forms so ATP has no access to remove chain terminators (*127*) (Figure 3.1). This model answered some important questions but failed to solve other questions. For example, why does the TAMs mutant selectively excise AZTMP with a much higher rate than other NRTIs? A possible explanation is that the steric

hindrance from azido group interferes with the binding of the next complementary nucleotide, which alleviates its inhibitory effect and allows an AZTMP terminated primer to more readily access the N site (*59*). However, not all 3'-azido analogs are good substrates for excision (*80*) and excision efficiencies of many other NRTIs are still lower than that of AZTMP even without the inhibition from an incoming nucleotide (*141*), which suggests that other components of NRTIs, such as bases, also contribute to the selectivity of excision. How the TAMs mutations confer enhanced excision is another unresolved question. Although some studies claimed that the main role of the TAMs was to increase the ability of RT to bind ATP (*76, 115*), kinetic analyses reported comparable binding affinities of ATP in the presence of TAMs mutations (*78*). Since the equilibrium of translocation between N site and P site is the key point of the Boyer model, it's worth exploring if the TAMs mutations have any direct influence on the translocation status of HIVRT.

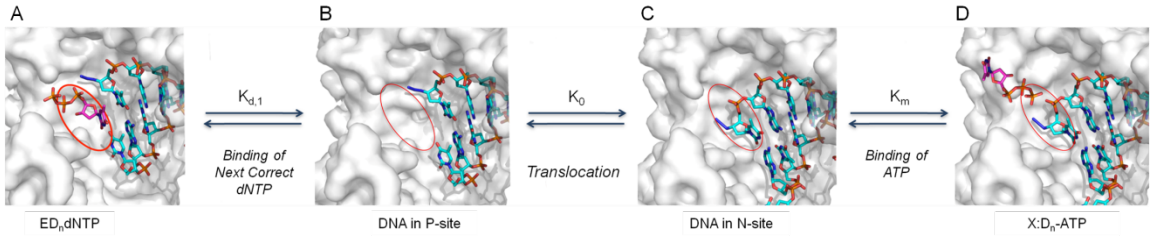



**Figure 3.1** *Competition between DEC Formation and Excision.* The RT structure is shown as grey surface. Primer/template is shown as blue sticks. The red ellipse denotes the incoming nucleotide binding site (N site). (A) Structure of HIVRT with a dead-end-complex (DEC) formed at the active site, where the N site is occupied by the next correct nucleotide (dCTP) so that ATP has no access for excision. (B) AZT terminated primer/template resides in the P site, allowing dCTP to bind to the N site. (C) AZT terminated primer/template occupies N site blocking nucleotide binding. (D) ATP binds and is ready to initiate the unblocking of the AZT-terminated primer. Drawn using Pymol from PDB: 1N6Q and 1N5Y (*59*).

An additional T69S mutation and a nearby dipeptide insertion (usually SS, SA or SG) in the TAMs backbone give rise to enhanced breadth and efficiency of NRTI excision (*117-121*). The fingers insertion can destabilize the formation of ternary dead-end complexes so that inhibition against the excision of these NRTIs would be reduced in the presence of physiological concentrations of dNTPs (*115, 117, 120*). However, our previous studies showed that introduction of a Ser-Ser insertion into the TAMs background actually decreased the excision efficiency of AZTMP by ~2.5-fold while weakening the binding of the next correct nucleotide (dCTP) by 2-fold (Chapter 2). Unfortunately, the binding of dCTP is too weak to allow the 2-fold change to make any significant contribution to excision rates at physiological conditions.

In this chapter, studies are continued to explore the excision of another two thymidine analogs, d4T and ddT, by TAMs and 69 insertion complex mutants to explore how fingers insertions would broaden excision specificity. To address this question, primer unblocking and extension assays were performed to explore the concentration dependence of the reaction time course. Stopped flow and equilibrium titration assays were also conducted to accurately define the relevant kinetic parameters including the binding affinity for the next cognate nucleotide. In addition, comparisons of the excision of the three thymidine analogs by RT mutants shed some light on how drug components contribute to excision selectivity.

Although the reactions contributing to NRTI resistance have been described previously, there has been no accurate quantitative analysis to provide rigorous assessment of the relative contributions of each reaction to the net resistance. Assays for

rescue of an analog-terminated primer are complex in that they are dependent on five reactions: translocation of the primer to the N-site on the polymerase, ATP binding and excision of the analog-terminated primer, elongation of the primer by the addition of TTP and then again by dCTP, but the process is also inhibited by the binding of dCTP to the analog-terminated primer to form a dead-end complex when it is in the P site. The interplay between the various competing reactions and their additive effects on the net observable parameters is difficult to resolve by conventional analysis. Here we resolve this complexity by taking advantage of global data fitting to reveal the intrinsic rate and equilibrium constants governing each step leading to resistance to chain terminators.

## 3.2 MATERIALS AND METHODS

### Expression and Purification of Proteins Used in This Study

The method for the preparation of MDCC-HIVRT used in this study was described previously (*77, 86*). Briefly, the thymidine analogues mutant (TAMs) was produced by introducing six mutations (M41L, D67N, K70R, L210W, T215Y and k219E) in genes of both subunits using a Quikchange Multi kit (Stratagene). Based on genes of the TAMs mutant, the TAMs_T69SSSVD mutant was produced by introducing T69SSS followed by the introduction of A62V and a reverse mutation of N67D (Table 2.1). Both subunits for each mutant were individually expressed in T7 Express Competent *E.coli* (New England Biolabs; NEB), mixed in a 1:1 ratio and purified by tandem Q-Sepharose and Bio-Rex 70 columns followed by a single-strand DNA affinity column. Following MDCC labeling, excess MDCC was removed using a Bio-Rex 70 column. The labeled

enzymes were assayed by pre-steady-state burst experiments to determine the active site concentrations, divided into aliquots, and stored at -80°C.

**DNA Substrates for Kinetic Studies**

Primer DNA terminated with ddTMP or d4TMP was made through enzymatic primer extension using RT (Table 3.1), as described in Chapter 2. Following the synthetic reactions, terminated 26mer primers (26ddTMP or 26d4TMP) were purified by gel extraction. Annealing of the 26mer primers and 45mer template was carried out by mixing the two oligonucleotides in a 1:1 molar ratio at 95°C for 5min, followed by slow cooling to room temperature. Primers used in quench flow assays were 5'-$^{32}$P-labeled using T4 polynucleotide kinase (NEB).

**Table 3.1** *DNA substrates for excision assays*

*26ddT/45-A:*

25mer: 5′-GCCTCGCAGCCGTCCAACCAACTCA$T_{dd}$-3′
45mer: 3′-CGGAGCGTCGGCAGGTTGGTTGAGT**A**GCAGCTAGGTTACGGCAGG-5′

*26d4T/45-A:*

25mer: 5′-GCCTCGCAGCCGTCCAACCAACTCA$T_{4}$-3′
45mer: 3′-CGGAGCGTCGGCAGGTTGGTTGAGT**A**GCAGCTAGGTTACGGCAGG-5′

**Quench Flow Kinetic Assays**

Excision of ddTMP or d4TMP from the 3' end of blocked primers and concomitant primer extension were measured using a hand-quench assay. The 26/45mer primer-template (100 nM) was pre-incubated with either TAMs or TAMs_T69SSSVD mutants (150 nM) at 4°C for 10 min in a buffer containing 50 mM Tris pH7.5, 100 mM

KCl and 10 mM Magnesium acetate (all concentrations final). Primer rescue and extension were initiated by adding 100 μM TTP and 1 μM dCTP in the presence of various concentrations of ATP and 0.6 μM pyrophosphatase. A 10 μL reaction aliquot was mixed with 10 μL 0.5 M EDTA and 20 μL loading buffer to stop the reaction for each time point. Products were separated by 15% denaturing PAGE and quantified using a Typhoon scanner in ImageQuant 6.0 (GE). The experimental setup was similar for assays designed to assess the inhibitory effect of dCTP (the next correct nucleotide binding to the 3'end of a ddTMP or d4TMP-terminated primer) except that the concentration of ATP was fixed at 4.8 mM while various concentrations of dCTP were used.

**Stopped Flow Kinetic Assays**

Binding affinities of dCTP to ddTMP or d4TMP terminated primers were measured by fluorescence stopped-flow and equilibrium titration assays using MDCC-labeled RT (*77, 86*). The time dependence of MDCC fluorescence changes corresponding to dCTP-induced enzyme conformational change was monitored using an Auto SF-200x (www.kintekcorp.com). A pre-formed enzyme-DNA duplex (100 nM of MDCC-labeled HIV RT mutants and 150 nM 26ddTMP or 26d4TMP/45mer DNA) was rapidly mixed with various concentrations of dCTP in the stopped flow for 10 or 60 s, depending on the rate the reaction. Time was shown in logarithmic scale to illustrate the fast and the slower phases. MDCC fluorescence was excited at 425 nm and fluorescence was monitored at 475 nm with a 50 nm band-pass filter (Semrock). To get more information of dCTP binding to thymidine analog-terminated primers, equilibrium titration assays were also

performed using a KinTek TMX titration module (www.kintekcorp.com) (*87*). Briefly, the preformed enzyme-DNA duplexes (100 nM of MDCC-labeled HIV RT mutants and 150 nM 26ddTMP or 26d4TMP/45mer DNA) were continuously mixed with increasing concentrations of dCTP while recording the fluorescent intensities at equilibrium. Different final concentrations (15-400 μM) were chosen based on varying binding affinities of dCTP to the enzyme-DNA duplexes.

**Data Analysis**

A representative data set (data obtained from the TAMs/26ddTMP duplex) were analyzed by both conventional and global fitting to demonstrate how the combination of the two measurements allows resolution of binding and dissociation rates: on the one hand, conventional fitting provide numbers we could use as first estimations in the global analysis and they provide a the rationale for the models used in the more accurate global fitting. On the other hand, conventional fitting analyzes data separately and can ignore valuable linkage among experiments while global analysis treats the data as a whole and fits them to a single model, and therefore yields more accurate results and can reveal additional kinetic details.

**Conventional Data Fitting**

The time course of reactions were fit to single (Figure 3.2A and C) or double (Figure 3.2E) exponential functions using GraFit 5 (Erithacus Software Limited) (*86*). The ATP concentration dependence of the observed rates (Figure 3.2B) was fit to a hyperbola (Equation 3.1) to define the rate of excision, $k_{exc}$, and $K_m$ value for ATP.

$$k = \frac{k_{exc}[\mathrm{ATP}]}{K_m + [\mathrm{ATP}]} \qquad \text{(Equation 3.1)}$$

Similarly, dCTP concentration dependence of its inhibitory effect was fit to Equation 3.2 to derive the apparent $K_d$ of dCTP, where $v_0$ is the initial excision rate without any inhibition from dCTP (Figure 3.2D).

$$v = v_0 \times (1 - \frac{[\mathrm{dCTP}]}{K_{d,app} + [\mathrm{dCTP}]}) \qquad \text{(Equation 3.2)}$$

Apparent binding affinity of dCTP was also determined from the hyperbolic analysis (Equation 3.3) of dCTP concentration dependence of amplitude changes in stopped flow assays (Figure 3.2F), where $A_{max}$ represents the maximum amplitude of fluorescence decrease induced by the binding of dCTP.

$$\frac{A}{A_{max}} = \frac{[\mathrm{dCTP}]}{K_{d,app} + [\mathrm{dCTP}]} \qquad \text{(Equation 3.3)}$$

Equilibrium binding data were also fitted to a hyperbola (Equation 3.4), where $F_0$ is the starting fluorescence and $\Delta F$ is the change.

$$F = F_0 + \Delta F \frac{[\mathrm{dCTP}]}{K_{d,app} + [\mathrm{dCTP}]} \qquad \text{(Equation 3.4)}$$

**Global Data Fitting**

In order to derive kinetic constants governing the competition between chain terminator removal and dead-end complex formation, the data were first globally fit to a model containing a three-step nucleotide binding mechanism, and a key translocation step between N site and P site. As shown in Scheme 3.1, $ED_{25}ddT$ represents an enzyme-DNA

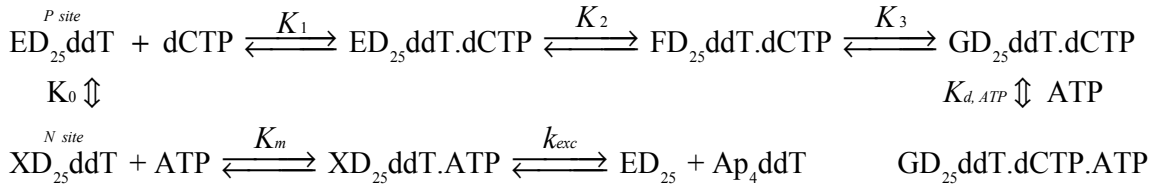
complex with DNA terminated by ddT, FD and GD represent closed enzyme states with altered fluorescence, XD represents the enzyme state with DNA bound to the N site and $Ap_4ddT$ is the dinucleotide tetraphosphate product of the excision reaction.

**Scheme 3.1** *Competition between DEC Formation (3-step) and ddTMP Removal*

$$\begin{array}{ccccccc} {}^{P\ site}ED_{25}ddT & + & dCTP & \overset{K_1}{\rightleftharpoons} & ED_{25}ddT.dCTP & \overset{K_2}{\rightleftharpoons} & FD_{25}ddT.dCTP \overset{K_3}{\rightleftharpoons} GD_{25}ddT.dCTP \\ K_0 \Updownarrow & & & & & & \\ {}^{N\ site}XD_{25}ddT & + & ATP & \overset{K_m}{\rightleftharpoons} & XD_{25}ddT.ATP & \overset{k_{exc}}{\rightleftharpoons} & ED_{25} + Ap_4ddT \end{array}$$

Excision efficiencies were determined by examining the ATP concentration dependence of the rate of rescue while the inhibitory effect of dCTP was quantified through similar assays with a fixed ATP concentration and various dCTP concentrations. The stopped flow data further defined the binding affinity of dCTP to the 3' end of analog-terminated primers in the absence of ATP. The smooth curves in each figure panel represent the best fit of all of the data according to the model (Figure 3.3). In order to better understand the apparent tighter binding of dCTP defined by the quench flow data, the same set of data were also fit to an improved model (Scheme 3.2), in which ATP can also bind to the $GD_{25}ddT.dCTP$ (or $GD_{25}d4T.dCTP$) species to explain the observed tighter apparent dCTP binding in the presence of ATP (Figure 3.5).

**Scheme 3.2** *Competition between DEC Formation (4-step) and ddTMP Removal*




The remaining data sets (Figure 3.6-3.8) were also fit to this improved model since it described the reactions more accurately. Parameters derived from global fitting are summarized in Table 3.2 and 3.3. In Table 3.3, the net dissociation constant of dCTP in the absence of ATP ($K_{d,dCTP}$) is given by the formula: $K_{d,dCTP} = 1 / (K_1 \times (1 + K_2 + K_2K_3))$, where $K_1$, $K_2$ and $K_3$ are equilibrium constants governing the ground-state binding of dCTP, and two dCTP-induced conformational change steps (Scheme 3.2). ATP binding ($K_{d,ATP}$) induced an additional conformational change and resulted in the apparently tighter binding of dCTP ($K_{dCTP,app}$), which is defined by equation $K_{dCTP,app} = K_{d,dCTP} / (1 + [ATP]/K_{d,ATP})$, where $K_{d,dCTP}$ was just defined and $K_{d,ATP}$ is the dissociation constant governing the binding of ATP to the $GD_{25}$ddT.dCTP species. The equilibrium constant for the translocation step ($K_0$, defined in the reverse direction of translocation during processive polymerization) and binding affinity of ATP to $XD_{25}$ddT ($K_{m,ATP}$) together determined the fraction of E-$DNA_{dd}$ duplex residing at the P site that is able to bind dCTP and therefore should also be included when calculating the overall apparent $K_d$ of dCTP ($K_{d,app}$): $K_{d,app} = K_{dCTP,app} \times (1 + K_0 \times (1 + [ATP]/K_{m,ATP}))$. The complete derivations of these formulas can be found below. In the absence of ATP, the overall apparent $K_d$ of dCTP can be easily determined as $K_{d,dCTP} \times (1 + K_0)$.

*Derivation of $K_{d,dCTP}$:*

$$K_1 = \frac{[ED_{25}ddT.dCTP]}{[ED_{25}ddT]\times[dCTP]}\ , \quad [ED_{25}ddT.dCTP] = K_1 \times [ED_{25}ddT]\times[dCTP]$$

$$K_2 = \frac{[FD_{25}ddT.dCTP]}{[ED_{25}ddT.dCTP]}\ , \qquad [FD_{25}ddT.dCTP] = K_2 \times [ED_{25}ddT.dCTP] = K_1 \times K_2 \times [ED_{25}ddT]\times[dCTP]$$

$$K_3 = \frac{[GD_{25}ddT.dCTP]}{[FD_{25}ddT.dCTP]}\ , \qquad [GD_{25}ddT.dCTP] = K_3 \times [FD_{25}ddT.dCTP] = K_1 \times K_2 \times K_3 \times [ED_{25}ddT]\times[dCTP]$$

$$\theta = \frac{[ED_{25}ddT.dCTP] + [FD_{25}ddT.dCTP] + [GD_{25}ddT.dCTP]}{[ED_{25}ddT] + [ED_{25}ddT.dCTP] + [FD_{25}ddT.dCTP] + [GD_{25}ddT.dCTP]}$$

$$= \frac{K_1 \times [ED_{25}ddT]\times[dCTP]\times(1 + K_2 + K_2 \times K_3)}{[ED_{25}ddT] + K_1 \times [ED_{25}ddT]\times[dCTP]\times(1 + K_2 + K_2 \times K_3)}$$

$$= \frac{K_1 \times [dCTP]\times(1 + K_2 + K_2 \times K_3)}{1 + K_1 \times [dCTP]\times(1 + K_2 + K_2 \times K_3)}$$

$$= \frac{[dCTP]}{\frac{1}{K_1 \times (1 + K_2 + K_2 \times K_3)} + [dCTP]} = \frac{[dCTP]}{\frac{1}{K_1 \times (1 + K_2 \times (1 + K_3))} + [dCTP]}$$

*Derivation of $K_{dCTP,app}$:*

To make the derivation process simpler and more clear, we simplify the steps before ATP binding to only one step, which includes ground state binding and two conformational change steps (showing in Scheme 3.2) and which are governed by the net dissociation constant of dCTP, that is, $K_{d,dCTP}$.

So we have:

$$K_{d,dCTP} = \frac{[ED_{25}ddT]\times[dCTP]}{[GD_{25}ddT.dCTP]}, \quad [GD_{25}ddT.dCTP] = \frac{[ED_{25}ddT]\times[dCTP]}{K_{d,dCTP}}$$

$$K_{d,ATP} = \frac{[GD_{25}ddT.dCTP]\times[ATP]}{[GD_{25}ddT.dCTP.ATP]}, \quad [GD_{25}ddT.dCTP.ATP] = \frac{[GD_{25}ddT.dCTP]\times[ATP]}{K_{d,ATP}}$$

$$= \frac{[ED_{25}ddT]\times[dCTP]\times[ATP]}{K_{d,dCTP}\times K_{d,ATP}}$$

$$\theta = \frac{[GD_{25}ddT.dCTP] + [GD_{25}ddT.dCTP.ATP]}{[ED_{25}ddT] + [GD_{25}ddT.dCTP] + [GD_{25}ddT.dCTP.ATP]}$$

$$= \frac{\frac{[ED_{25}ddT]\times[dCTP]}{K_{d,dCTP}} + \frac{[ED_{25}ddT]\times[dCTP]\times[ATP]}{K_{d,dCTP}\times K_{d,ATP}}}{[ED_{25}ddT] + \frac{[ED_{25}ddT]\times[dCTP]}{K_{d,dCTP}} + \frac{[ED_{25}ddT]\times[dCTP]}{K_{d,dCTP}\times K_{d,ATP}}} = \frac{\frac{[dCTP]}{K_{d,dCTP}} + \frac{[dCTP]\times[ATP]}{K_{d,dCTP}\times K_{d,ATP}}}{1 + \frac{[dCTP]}{K_{d,dCTP}} + \frac{[dCTP]\times[ATP]}{K_{d,dCTP}\times K_{d,ATP}}}$$

$$= \frac{\frac{[dCTP]}{K_{d,dCTP}} \times (1 + \frac{[ATP]}{K_{d,ATP}})}{1 + \frac{[dCTP]}{K_{d,dCTP}} \times (1 + \frac{[ATP]}{K_{d,ATP}})} = \frac{[dCTP]}{\frac{K_{d,dCTP}}{(1 + \frac{[ATP]}{K_{d,ATP}})} + [dCTP]}$$

*Derivation of $K_{d,app}$:*

To make the derivation more concise and straightforward, we again consider the binding of dCTP to E-$DNA_{dd}$ duplex before ATP binding as one step which is governed by $K_{d,dCTP}$. With this in mind, we have:

$$K_0 = \frac{[XD_{25}ddT]}{[ED_{25}ddT]}, \qquad [XD_{25}ddT] = K_0 \times [ED_{25}ddT]$$

$$K_{d,dCTP} = \frac{[ED_{25}ddT]\times[dCTP]}{[GD_{25}ddT.dCTP]}, \qquad [GD_{25}ddT.dCTP] = \frac{[ED_{25}ddT]\times[dCTP]}{K_{d,dCTP}}$$

$$K_{d,ATP} = \frac{[GD_{25}ddT.dCTP]\times[ATP]}{[GD_{25}ddT.dCTP.ATP]}, \quad [GD_{25}ddT.dCTP.ATP] = \frac{[GD_{25}ddT.dCTP]\times[ATP]}{K_{d,ATP}}$$

$$= \frac{[ED_{25}ddT]\times[dCTP]\times[ATP]}{K_{d,dCTP}\times K_{d,ATP}}$$

Since the excision rate is extremely slow relative to that of ATP dissociation, the binding of ATP comes to equilibrium so that $K_{m,ATP}$ reflects the true $K_d$ of ATP. In this case, we have:

$$K_m = \frac{[XD_{25}ddT]\times[ATP]}{[XD_{25}ddT.ATP]}, \quad [XD_{25}ddT.ATP] = \frac{[XD_{25}ddT]\times[ATP]}{K_m} = \frac{K_0\times[ED_{25}ddT]\times[ATP]}{K_m}, \text{ so}$$

$$\theta = \frac{[GD_{25}ddT.dCTP] + [GD_{25}ddT.dCTP.ATP]}{[ED_{25}ddT] + [GD_{25}ddT.dCTP] + [GD_{25}ddT.dCTP.ATP] + [XD_{25}ddT] + [XD_{25}ddT.ATP]}$$

$$= \frac{\frac{[ED_{25}ddT]\times[dCTP]}{K_{d,dCTP}} + \frac{[ED_{25}ddT]\times[dCTP]\times[ATP]}{K_{d,dCTP}\times K_{d,ATP}}}{[ED_{25}ddT] + \frac{[ED_{25}ddT]\times[dCTP]}{K_{d,dCTP}} + \frac{[ED_{25}ddT]\times[dCTP]\times[ATP]}{K_{d,dCTP}\times K_{d,ATP}} + K_0\times[ED_{25}ddT] + \frac{K_0\times[ED_{25}ddT]\times[ATP]}{K_m}}$$

$$= \frac{\frac{[dCTP]}{K_{d,dCTP}}\times(1+\frac{[ATP]}{K_{d,ATP}})}{1+\frac{[dCTP]}{K_{d,dCTP}}\times(1+\frac{[ATP]}{K_{d,ATP}}) + K_0\times(1+\frac{[ATP]}{K_m})}$$

Let $\theta = 0.5$, we have:

$$\frac{\frac{[dCTP]_{50}}{K_{d,dCTP}} \times (1 + \frac{[ATP]}{K_{d,ATP}})}{1 + \frac{[dCTP]_{50}}{K_{d,dCTP}} \times (1 + \frac{[ATP]}{K_{d,ATP}}) + K_0 \times (1 + \frac{[ATP]}{K_m})} = 0.5 ,$$

then we get:

$$\frac{[dCTP]_{50}}{K_{d,dCTP}} \times (1 + \frac{[ATP]}{K_{d,ATP}}) = 1 + K_0 \times (1 + \frac{[ATP]}{K_m})$$

Finally, we have:

$$[dCTP]_{50} = \frac{K_{d,dCTP} \times (1 + K_0 \times (1 + \frac{[ATP]}{K_m}))}{(1 + \frac{[ATP]}{K_{d,ATP}})} = \frac{K_{d,dCTP}}{(1 + \frac{[ATP]}{K_{d,ATP}})} \times (1 + K_0 \times (1 + \frac{[ATP]}{K_m}))$$

$$= K_{dCTP,app} \times (1 + K_0 \times (1 + \frac{[ATP]}{K_m})) = K_{d,app}$$

## 3.3 RESULTS

### Excision of ddTMP by the HIVRT_TAMs-MDCC

ATP dependent excision of AZTMP from the 3'end of the blocked primer has been identified as a significant source of resistance to AZT (*122*). In this study, the ATP concentration dependence of excision and extension were examined to determine excision efficiencies ($k_{exc}/K_{m,ATP}$) of two thymidine analogs (ddT and d4T) by TAMs mutants with or without the T69SSS fingers insertion. Since AZTMP-terminated primer rescue can be specifically inhibited by the next complementary nucleotide (*126-127*), dCTP (in our studies) concentration dependence of the excision rates of ddTMP or d4TMP by the RT mutants was determined to explore whether the fingers insertion has any effect on the ability of dCTP to inhibit excision. Stopped flow and equilibrium titration assays were also performed to further determine affinities for dCTP binding to analog-terminated primers.

ATP concentration dependent excision assays were first performed to determine the excision efficiency of ddTMP by the HIVRT_TAMs-MDCC (Figure 3.2A). The data were fit to a single exponential function at each ATP concentration and the concentration dependence of the observed rates was then fit to a hyperbola (Equation 1) to define the maximum excision rate, 0.0052 ± 0.0004 $s^{-1}$, and $K_m$ of ATP, 2.8 ± 0.4 mM, based on which, an excision efficiency of 1.9 ± 0.3 $M^{-1}s^{-1}$ was obtained (Figure 3.2B). Excision assays were also performed with a fixed concentration of ATP (4.8 mM) and various concentrations of dCTP (Figure 3.2C). Observed rates obtained by fitting the data to a single exponential function were fit to Equation 3.2, which yielded the apparent

dissociation constant of dCTP as 8.6 ± 0.6 μM (Figure 3.2D). Figure 3.2E shows the time dependence of the fluorescence changes measured by stopped flow methods during dCTP binding at several nucleotide concentrations. Since the fluorescence decreases were biphasic, the data were fit to a double exponential function. Amplitudes of the fluorescent signals were fit to Equation 3.3, from which an apparent $K_d$ of 41 ± 2 μM was obtained (Figure 3.2F). Figure 3.2G shows the measurement of the binding affinity of dCTP by the equilibrium titration assay, where the black line represents the hyperbolic analysis (Equation 3.4) of the data, which defines an apparent dissociation constant of 35 ± 0.1 μM. It is worth noting that $K_d$ values of dCTP obtained from fluorescence measurements are in good agreement with each other, which, however, are 4~5 fold higher than the number obtained by quench flow assays of rescue in the presence of ATP.

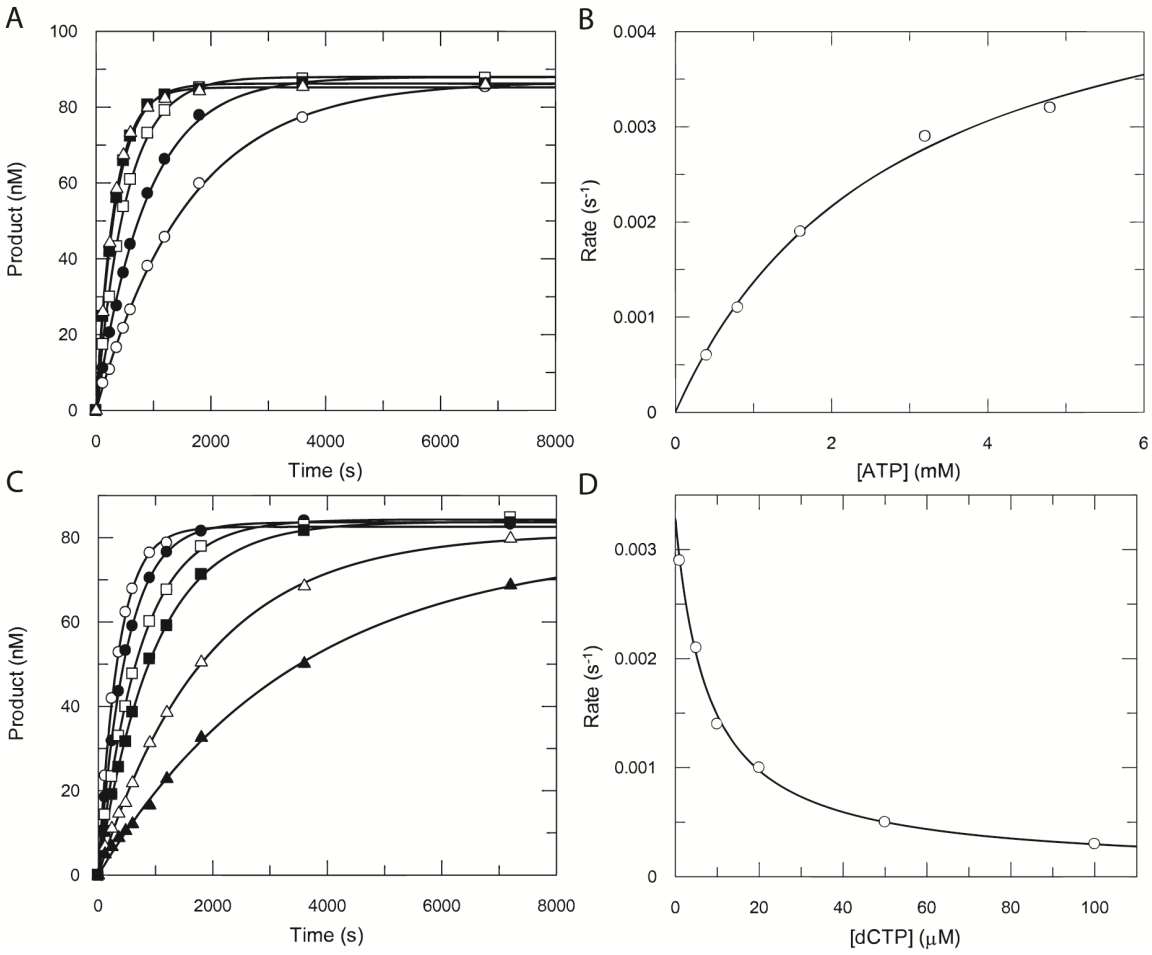



**Figure 3.2** (A-D) *Conventional fitting of ATP and dCTP concentration dependence of ddTMP excision by MDCC-labeled HIV RT_TAMs, as well as binding affinity of dCTP to preformed TAMs-E.DNA$_{dd}$ duplex (A-D).* ATP (A) and dCTP (C) concentration dependence of ddTMP excision and rescue as monitored by the formation of elongated product. A preformed TAMs/26ddTMP complex (150 nM MDCC-labeled HIVRT_ TAMs and 100 nM of 26ddTMP/45mer) was rapidly mixed with 100 μM TTP and 1 μM dCTP in the presence of various concentrations of ATP (0.4, 0.8, 1.6, 3.2, and 4.8 mM). Rates were obtained by nonlinear regression to single exponential function (A) and the concentration dependence of the observed rates was then fit to a hyperbola (Equation 3.1) to define the maximum excision rate ($k_{exc}$ = 0.005 ± 0.0004 $s^{-1}$), $K_m$ of ATP (2.8 ± 0.4 mM), and the specificity constant ($k_{exc}/K_{m,ATP}$) as 1.8 ± 0.3 $M^{-1}s^{-1}$ (B). (C) As to dCTP concentration dependent assays, the preformed enzyme-DNA duplex (150 nM MDCC-labeled HIVRT_TAMs and 100 nM of 26ddTMP/45mer) was rapidly mixed with 100 μM TTP and 4.8 mM ATP in the presence of various concentrations of dCTP (1, 5, 10, 20, 50 and 100 μM). Similarly, the concentration dependence of inhibition by dCTP was fit to a hyperbola (Equation 3.2) to define the apparent $K_d$ of dCTP as 8.6 ± 0.6 μM (D).

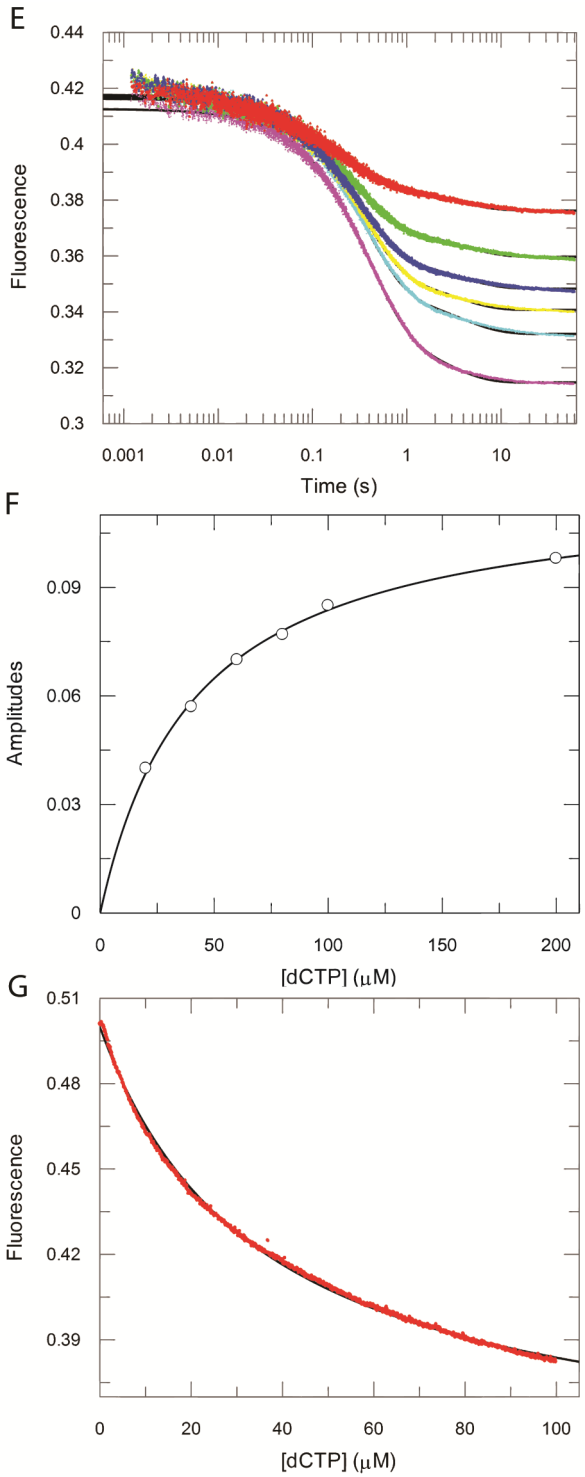



**Figure 3.2** (E-G) *Conventional fitting of ATP and dCTP concentration dependence of ddTMP excision by MDCC-labeled HIV RT_TAMs, as well as binding affinity of dCTP to preformed TAMs-E.DNA$_{dd}$ duplex (E-G).* Binding affinity of dCTP to enzyme-DNA complex terminated with ddTMP was measured by stopped flow (E), as well as equilibrium titration assays (G). (E) A preformed TAMs/ 26ddTMP complex (100 nM TAMs and 150 nM of 26ddTMP/45mer) was rapidly mixed with various concentrations of dCTP (20, 40, 60, 80, 100 and 200 μM). Fluorescence was observed by excitation of MDCC at 425 nm and monitoring emission with a 475 nm band-pass filter with a 50 nm bandwidth for 60 s (time was showed in logarithmic scale). Stopped flow data were fit to a double exponential function at each nucleotide concentration and the concentration dependence of the derived amplitudes ($A_1$+$A_2$) was then fit to a hyperbola (Equation 3.3) to define an apparent dissociation constant of dCTP as 41 ± 2 μM (F). (G) In the titration assay, the preformed TAMs/26ddTMP complex (100 nM TAMs and 150 nM of 26ddTMP/45mer) was continuously mixed with increasing concentrations of dCTP (0-100 μM) while recording fluorescence. Data were fit by nonlinear regression to a hyperbolic equation (Equation 3.4) to define an apparent dissociation constant of 35 ± 0.1 μM.

Global analysis was also performed for the four experiments, according to Scheme 3.1, as illustrated in *materials and methods* (Figure 3.3). This analysis demonstrates the simultaneous fitting of chemical quench flow data to monitor the rate of excision in the presence of various concentrations of ATP (Figure 3.3A), or dCTP (Figure 3.3B), the fluorescence data to monitor dCTP-induced conformational changes (Figure 3.3C), and the equilibrium binding affinity of dCTP (Figure 3.3D). The maximum excision rate and $K_m$ of ATP obtained from conventional fitting were set as starting estimates in the global fitting to derive other rate constants. However, fitting curves in Figure 3.3B were significantly discrepant from the real data, suggesting that the data cannot be accurately fit to the minimal model shown in Scheme 3.1. This was not so surprising since conventional fitting already showed the discrepancies among $K_d$ values of dCTP measured by quench flow and stopped flow methods, and this irreconcilable contradiction is more clearly demonstrated in the global analysis.

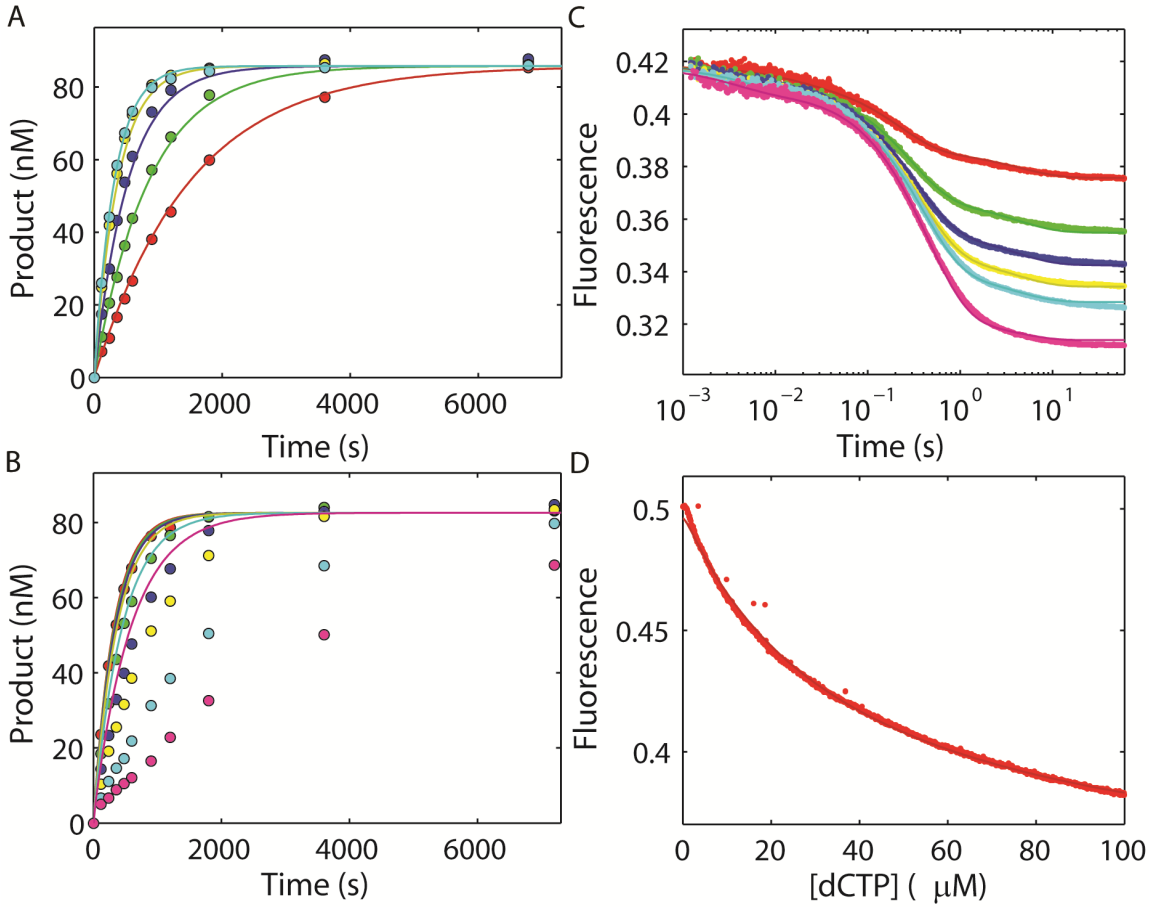



**Figure 3.3** *Global fitting of ddTMP excision and dCTP binding assays by MDCC-labeled HIVRT_TAMs.* ATP (A) and dCTP (B) concentration dependent excision/rescue assays, as well as dCTP binding (C) and equilibrium titration (D) assays shown in Figure 3.2 were also analyzed by global fitting of all the data to a single model (Scheme 3.1) using KinTek Explorer. Reaction time in panel C is shown in logarithmic scale. In each figure, the smooth lines show the best fit to the model. The fitting curves in panel B are significantly off from the real data since the discrepancies among binding affinities of dCTP derived from quench-flow data (B) and stopped-flow data (C/D) are significant, suggesting that important step(s) may be missing in the model.

Since the major difference between quench flow and stopped flow experiments is that binding affinity of dCTP was measured under the influence of high concentration of ATP (4.8 mM) in chemical quench assays, it is reasonable to speculate that ATP can bind

to the enzyme-$DNA_{dd}$-dCTP triplex and make the binding of dCTP tighter. This hypothesis was tested by mixing a pre-formed ternary complex either with buffer (Figure 3.4A) or in the presence of 4.8 mM ATP (Figure 3.4B) while recording the fluorescence. As was expected, when mixed with buffer, an increase of fluorescence was observed since dCTP dissociated from the ternary complex during the dilution and a new equilibrium was established. The rate was 2.1±0.1 $s^{-1}$, which is comparable to the dissociation rates of nucleotides from a closed ternary complex we previously measured (*77, 86*). Notably, in the presence of ATP, a decrease of fluorescence was observed with a rate of 0.45±0.01 $s^{-1}$, suggesting a slow conformational change after the binding of ATP.

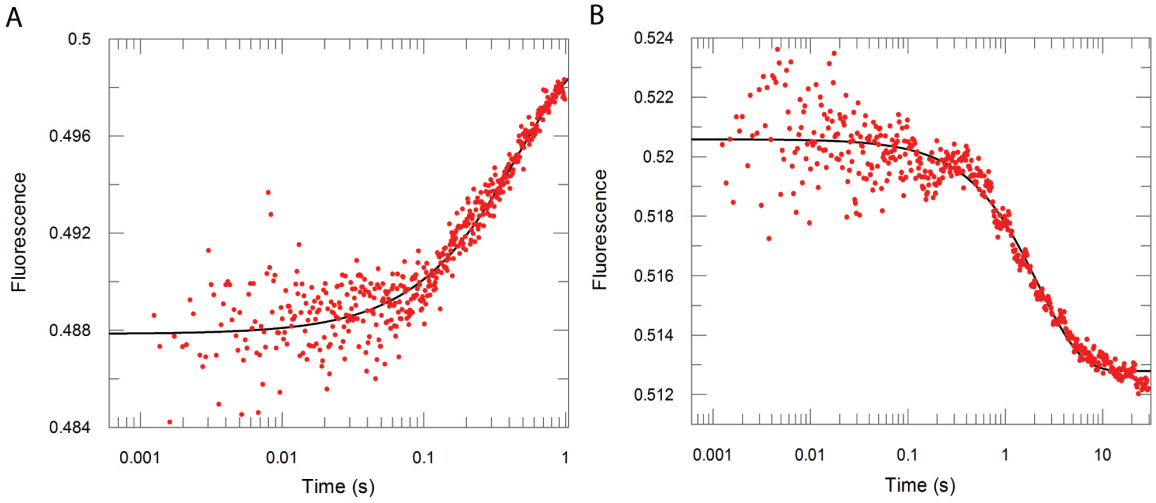



**Figure 3.4** *ATP makes the binding of dCTP tighter by inducing an additional isomerization step*. A pre-formed E-$DNA_{dd}$-dCTP ternary complex (100 nM MDCC-labeled HIVRT_TAMs, 150 nM of 26ddTMP/45mer, and 200 μM dCTP) was mixed with the reaction buffer (A) or 4.8 mM of ATP (B). Fluorescence was observed by excitation of MDCC at 425 nm and monitoring emission with a 475 nm band-pass filter with a 50 nm bandwidth for 1 (A) or 30 (B) s (time was showed in logarithmic scale). Stopped flow data were fit to a single exponential function. The rate of fluorescence increase (A) due to dCTP dissociation is 2.1±0.1 $s^{-1}$. While the rate of fluorescence decrease (B) induced by the ATP binding is 0.45±0.01 $s^{-1}$.

Theoretically, ATP can bind to the triplex during any stage of the binding event, either the ground-state binding ($ED_{25}$ddT.dCTP), or any one of the two conformational changes ($FD_{25}$ddT.dCTP or $GD_{25}$ddT.dCTP). In order to make the model as simple as possible, it was assumed that ATP only binds to the $GD_{25}$ddT.dCTP species. With this consideration in mind, the data were then globally fit to a modified model (Scheme 3.2) to derive rate constants that were used to calculate kinetic parameters governing the excision efficiency and binding of dCTP (Figure 3.5), as summarized in Table 3.2 and 3.3. The excision rate ($k_{exc}$) was 0.0047 ± 0.0003 $s^{-1}$, the $K_m$ of ATP was 2.6 ± 0.05 mM, yielding a $k_{exc}/K_{m,ATP}$ of 1.8 ± 0.1 $M^{-1}s^{-1}$ (Table 3.2), all of which agree with values for $k_{exc}$, $K_{m,ATP}$ and $k_{exc}/K_{m,ATP}$ obtained from conventional fitting. Global analysis resulted in a value of 3.3 ± 0.7 μM ($K_{d,dCTP} = 1/(K_1 \times (1 + K_2 + K_2K_3))$) governing the binding of dCTP to RTs locating at the P site in the absence of ATP. However, binding of ATP to $GD_{25}$ddT.dCTP ($K_{d,ATP}$ = 0.06 ± 0.003 mM) made the binding tighter by ~80 fold and defined the apparent $K_d$ of dCTP as 0.04 ± 0.007 μM ($K_{dCTP,app} = K_{d,dCTP} /(1 + [ATP]/K_{d,ATP})$). The ratio of the fluorescent amplitude for the fast phase to that of the slower phase defines an equilibrium constant ($K_0$) governing the reverse translocation between P site and N site (Figure 3.5C and Figure 3.10) of 12 ± 1. These values produced overall $K_d$ values for dCTP binding with ($K_{d,app} = K_{dCTP,app} \times (1 + K_0 \times (1 + [ATP]/K_{m,ATP}))$) or without ($K_{d,app} = K_{d,dCTP} \times (1 + K_0)$) ATP as 1.3 ± 0.3 and 41 ± 10 μM, respectively. The former is smaller than the $K_d$ of 8.6 ± 0.6 μM obtained from the conventional fitting of quench flow data (Figure 3.2D and Table 3.4) but the latter agrees

with $K_d$ values (41 ± 2 or 35 ± 0.1 μM) derived from fluorescence measurements (Figure 3.2F and G).

**Table 3.2** *Kinetic Parameters Governing Thymidine Analog Excision Efficiencies*

| | Kinetic Parameters | | | |
|---|---|---|---|---|
| | $k_{exc}$ ($s^{-1}$) | $K_{m,ATP}$ (mM) | $k_{exc}/K_{m,ATP}$ ($M^{-1}s^{-1}$) | Excision Rates ×1000 ($s^{-1}$, 50 μM dCTP) |
| TAMs / 26ddT | 0.0047±0.0003 | 2.6±0.05 | 1.8±0.1 | 0.5±0.02 |
| TAMs_T69SSSVD / 26ddT | 0.0032±0.0002 | 6.3±0.06 | 0.5±0.03 | 0.7±0.03 |
| TAMs / 26d4T | 0.022±0.0001 | 2.7±0.2 | 8.1±0.6 | 0.6±0.07 |
| TAMs_T69SSSVD / 26d4T | 0.014±0.0007 | 3.7±0.4 | 3.8±0.4 | 0.9±0.03 |

**Table 3.3** *Kinetic Parameters Governing the Binding of Next Correct Nucleotide*

| | Kinetic Parameters | | | | |
|---|---|---|---|---|---|
| | $K_{d,dCTP}$ (μM)[a] | $K_{d,ATP}$ (mM)[b] | $K_{dCTP,app}$ (μM)[c] | $K_0$ | $K_{d,app}$ (μM)[d] |
| TAMs / 26ddT | 3.3±0.7 | 0.06±0.003 | 0.04±0.007 | 12±1 | 1.3±0.3 |
| TAMs_T69SSSVD / 26ddT | 7.1±1.5 | 0.32±0.01 | 0.45±0.09 | 12±1 | 9.9±2.5 |
| TAMs / 26d4T | 0.6±0.05 | 0.27±0.01 | 0.03±0.003 | 2.8±0.1 | 0.26±0.04 |
| TAMs_T69SSSVD / 26d4T | 0.7±0.1 | 1.6±0.2 | 0.18±0.04 | 5.6±0.3 | 2.5±0.6 |

[a] Net dissociation constant of dCTP ($K_{d,dCTP}$) is defined by the product of equilibrium constants up to the catalytic step (in the absence of ATP): $K_{d,dCTP} = 1/(K_1 \times (1+K_2+K_2K_3))$.
[b] $K_{d,ATP}$ is the dissociation constant governing the binding of ATP to the $GD_{25}$ddT.dCTP.
[c] Apparent dissociation constant of dCTP in the presence of ATP is defined by the formula: $K_{dCTP,app} = K_{d,dCTP} / (1 + [ATP] / K_{d,ATP})$.
[d] Apparent dissociation constant is the product of the apparent dissociation constant of dCTP and the equilibrium constant of the translocation step from P site to N site: $K_{d,app} = K_{dCTP,app} \times (1 + K_0 \times (1+[ATP]/K_{m,ATP}))$.

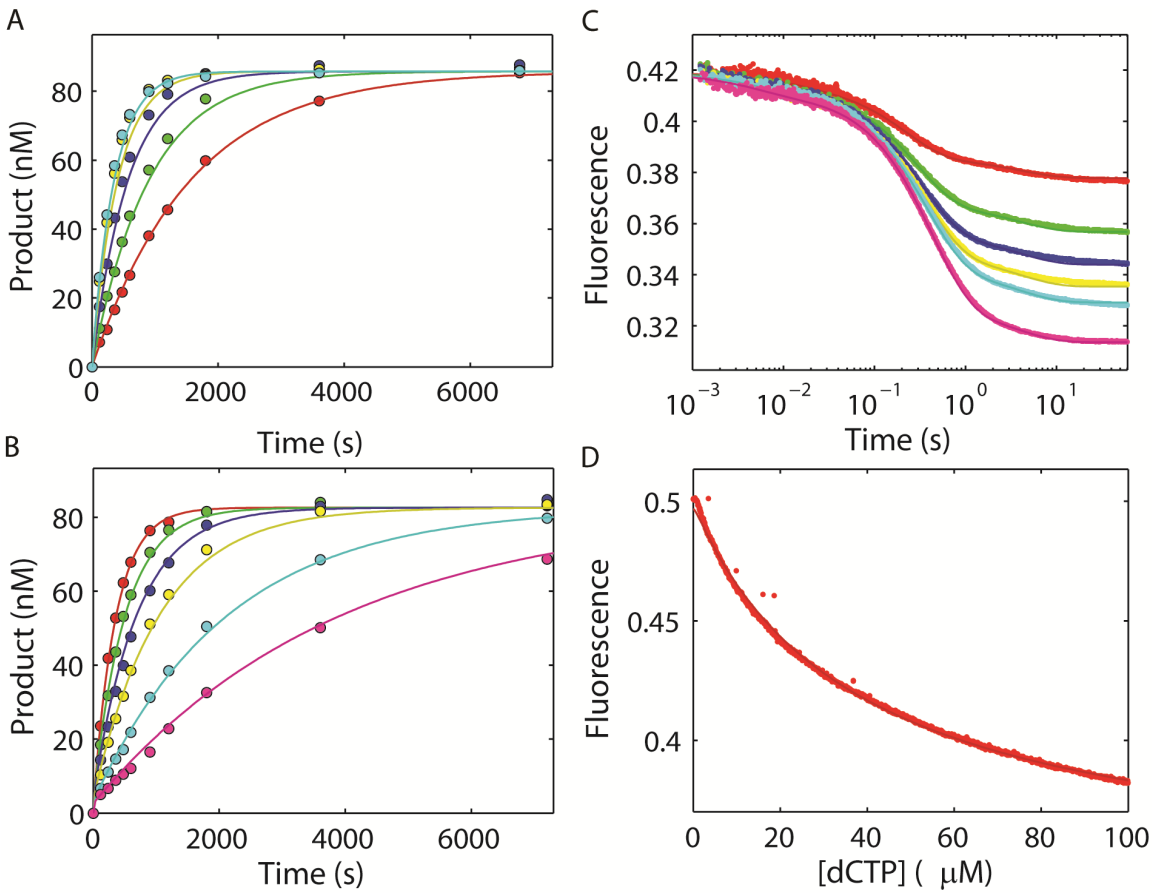



**Figure 3.5** *Global fitting of ddTMP excision and dCTP binding assays by MDCC-labeled HIVRT_TAMs to an improved model.* The same set of data were fit to an improved model shown in Scheme 3.2, in which binding of ATP onto the ternary complex "G-DNA$_{dd}$-dCTP" can induce the tighter binding of dCTP observed in the quench-flow data. All curves represent the best fit to get the rate constants shown in Table 3.2 and 3.3.

**Excision of ddTMP by the HIVRT_TAMs_T69SSSVD-MDCC**

Similar experiments and data analyses were performed to determine the kinetics governing ddTMP excision by and dCTP binding to the TAMs_T69SSSVD mutant, as shown in Figure 3.6 and summarized in Table 3.2 and 3.3. The rate of excision was 0.0032 ± 0.0002 $s^{-1}$, and the $K_m$ for ATP was 6.3 ± 0.06 mM, as defined by the chemical

quench flow data (Figure 3.6A). The fingers insertion reduced the excision efficiency by 3.6-fold, with a specificity constant of 0.5 ± 0.03 $M^{-1}s^{-1}$ (Table 3.2). The net dissociation constant of dCTP ($K_{d,dCTP}$) was 7.1 ± 1.5 μM. However, fingers insertion weakened the binding of ATP to $GD_{25}$ddT.dCTP by 5-fold ($K_{d,ATP}$ = 0.32 ± 0.01 mM), which finally made the binding of dCTP in the presence of ATP weaker by 11 fold ($K_{dCTP,app}$ = 0.45 ± 0.09 μM). Measurements of the equilibrium constant governing translocation ($K_0$ = N/P ratio of DNA binding) were fit to a value of 12 ± 1, suggesting that this mutant favors the binding of DNA at the N site where the drug is positioned to be excised efficiently. Taken these values altogether, the overall apparent $K_d$ for dCTP was determined as 9.9 ± 2.5 μM (Table 3.3), so the binding of dCTP to TAMs_T69SSSVD/26ddTMP was 7-fold weaker than to the TAMs/26ddTMP duplex, suggesting that the excision of ddTMP by the mutant with fingers insertion might be more resistant to the inhibition by dCTP binding to form the dead end complex.

**Excision of d4TMP by the HIVRT_TAMs-MDCC and HIVRT_TAMs_T69SSSVD-MDCC**

ATP concentration dependence of the rates of excision of d4TMP by the TAMs mutant defined a maximum rate of excision of 0.022 ± 0.0001 $s^{-1}$, a $K_m$ for ATP of 2.7 ± 0.2 mM, the excision efficiency was then calculated as $k_{exc}/K_{m,ATP}$ = 8.1 ± 0.6 $M^{-1}s^{-1}$, which is 4.5-fold higher than that observed for the excision of ddTMP (Figure 3.7A and Table 3.2). Data contained within the titration curve and families of curves for the dCTP concentration dependence of excision rates and fluorescence changes (Figure 3.7B-D) were used to define a net dissociation constant of dCTP ($K_{d,dCTP}$) as 0.6 ± 0.05 μM, a

dissociation constant governing the binding of ATP to the triplex ($K_{d,ATP}$) as 0.27 ± 0.01 mM, which leads to an apparent dissociation constant of dCTP ($K_{dCTP,app}$) of 0.03±0.003 μM (Table 3.3). The global analysis resulted in a $K_0$ of 2.8±0.1, suggesting that the TAMs mutant still favor binding the d4TMP-terminated primer/template at the N site, but with less than the ddTMP-terminated primer/template. In the presence of ATP, the binding of dCTP to TAMs/26d4TMP was 5-fold tighter than to the TAMs/26ddTMP duplex, with an overall apparent $K_d$ ($K_{d,app}$) of 0.26 ± 0.04 μM, suggesting that the excision of d4TMP by TAMs might be more susceptible to the inhibition by the next correct nucleotide (Table 3.3).

Excision of d4TMP by the TAMs_T69SSSVD mutant and the binding of dCTP to the E-$DNA_{d4T}$ duplex (Figure 3.8) were also globally fit to the four-step nucleotide binding model (Scheme 3.2). To determine the maximum rate of excision and $K_m$ for ATP, the concentration dependence was examined (Figure 3.8A) to obtain a rate of 0.014 ± 0.0007 $s^{-1}$, and a $K_{m,ATP}$ of 3.7 ± 0.4 mM, and the excision efficiency was determined to be 3.8 ± 0.4 $M^{-1}s^{-1}$, indicating that the fingers insertion decreased the excision efficiency by 2.2-fold (Table 3.2). Rate constants obtained from global analysis afforded a $K_{d,dCTP}$ = 0.7 ± 0.1 μM governing the binding of dCTP to the E.$DNA_{d4T}$ duplex located at the P site. Then the fingers insertion made the binding of ATP to the $GD_{25}$d4T.dCTP complex 6-fold weaker, with a $K_{d,ATP}$ value of 1.6 ± 0.2 mM. The apparent $K_d$ for dCTP in the presence of ATP ($K_{dCTP,app}$) was then determined as 0.18 ± 0.04 μM. The equilibrium constant governing translocation ($K_0$) was 5.6 ± 0.3 and the overall apparent $K_d$ for dCTP ($K_{d,app}$) was then calculated as 2.5 ± 0.6 μM, which is 10-fold larger than the value

obtained for the TAMs mutant (0.26 ± 0.04 μM). Again, our data supported that the fingers insertion weakened the binding of dCTP to a thymidine analog terminated primer and therefore might reduce its inhibition towards the excision. Global fitting results for ddTMP and d4TMP excision by TAMs and TAMs_T69SSSVD mutants are summarized in Tables 3.2 and 3.3.

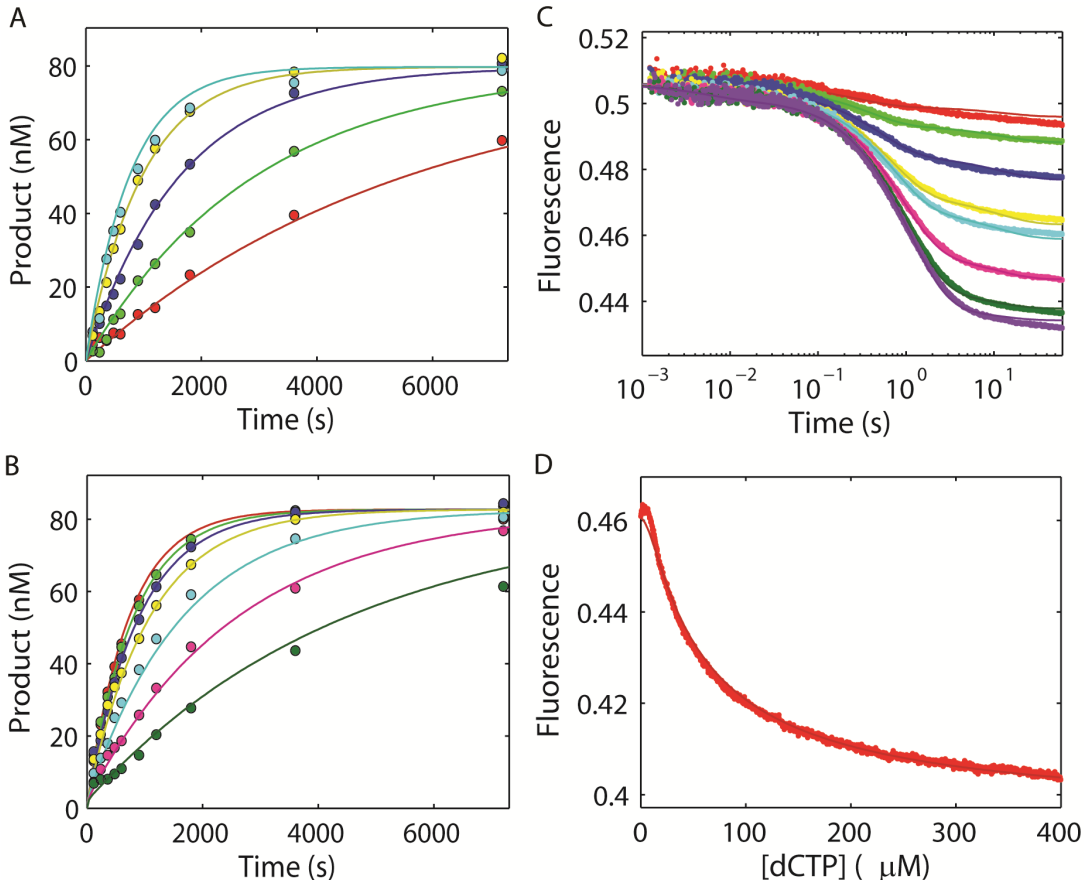



**Figure 3.6** *Global fitting of ddTMP excision and dCTP binding assays by MDCC-labeled HIVRT_TAMs_T69SSSVD.* ATP (A) and dCTP (B) concentration dependence of ddTMP excision and rescue as monitored by the formation of elongated product. (A) A preformed TAMs_T69SSSVD/26ddTMP complex (150 nM MDCC-labeled mutant and 100 nM of 26ddT/45mer) was mixed with 100 μM TTP and 1 μM dCTP in the presence of various concentrations of ATP (0.4, 0.8, 1.6, 3.2, and 4.8 mM). (B) A preformed enzyme-DNA duplex (150 nM MDCC-labeled RT mutant and 100nM of 26ddT/45mer) was mixed with 100 μM TTP and 4.8 mM ATP in the presence of various concentrations of dCTP (1, 5, 10, 20, 50, 100 and 200μM). Binding affinity of dCTP to enzyme-DNA complex terminated with ddTMP was measured by stopped-flow fluorescence (C), as well as equilibrium fluorescence titration assays (D). (C) A preformed TAMs_T69SSSVD /26ddTMP complex (100 nM mutant and 150 nM of 26ddT/45mer) was rapidly mixed with various concentrations of dCTP (10, 20, 40, 80, 100, 200, 400 and 600 μM). Fluorescence was observed for 60 s (time is showed on a logarithmic scale). (D) In the titration assay, the preformed TAMs_T69SSSVD/26ddTMP complex (100 nM mutant and 150nM of 26ddT/45mer) was continuously mixed with increasing concentrations of dCTP (0-400 μM) while recording fluorescence. All curves represent the best global fit to get the rate constants shown in Table 3.2 and 3.3.

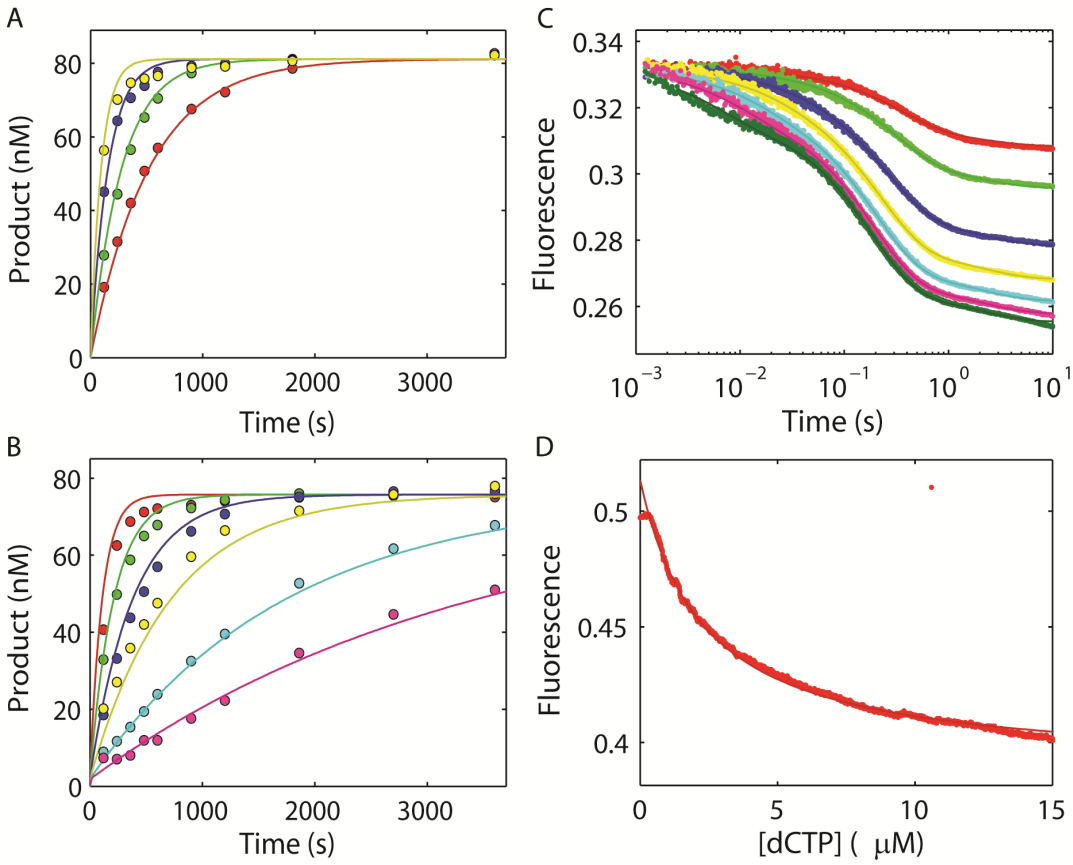



**Figure 3.7** *Global fitting of d4TMP excision and dCTP binding assays by MDCC-labeled HIVRT_TAMs.* ATP (A) and dCTP (B) concentration dependence of d4TMP excision and rescue as monitored by the formation of elongated product. (A) A preformed TAMs/26ddTMP complex (150 nM MDCC-labeled RT mutant and 100 nM of 26d4T/45mer) was rapidly mixed with 100 μM TTP and 1 μM dCTP in the presence of various concentrations of ATP (0.4, 0.8, 1.6, and 3.2 mM). (B) A preformed enzyme-DNA duplex (150 nM MDCC-labeled RT mutant and 100 nM of 26d4T/45mer) was rapidly mixed with 100 μM TTP and 4.8 mM ATP in the presence of various concentrations of dCTP (2, 5, 10, 20, 50, and 100 μM). Binding affinity of dCTP to enzyme-DNA complex terminated with d4TMP was measured by ensemble stopped-flow (C), as well as equilibrium titration assays (D). (C) A preformed TAMs/26d4TMP complex (100 nM TAMs and 150 nM of 26d4T/45mer) was rapidly mixed with various concentrations of dCTP (1, 2, 5, 10, 20, 40 and 100 μM). Fluorescence was observed for 10 s (time was showed in logarithmic scale). (D) In the titration assay, the preformed TAMs/26d4TMP complex (100 nM TAMs and 150 nM of 26d4T/45mer) was continuously mixed with increasing concentrations of dCTP (0-15 μM) while recording fluorescence. All curves represent the best fit to get the rate constants shown in Table 3.2 and 3.3.

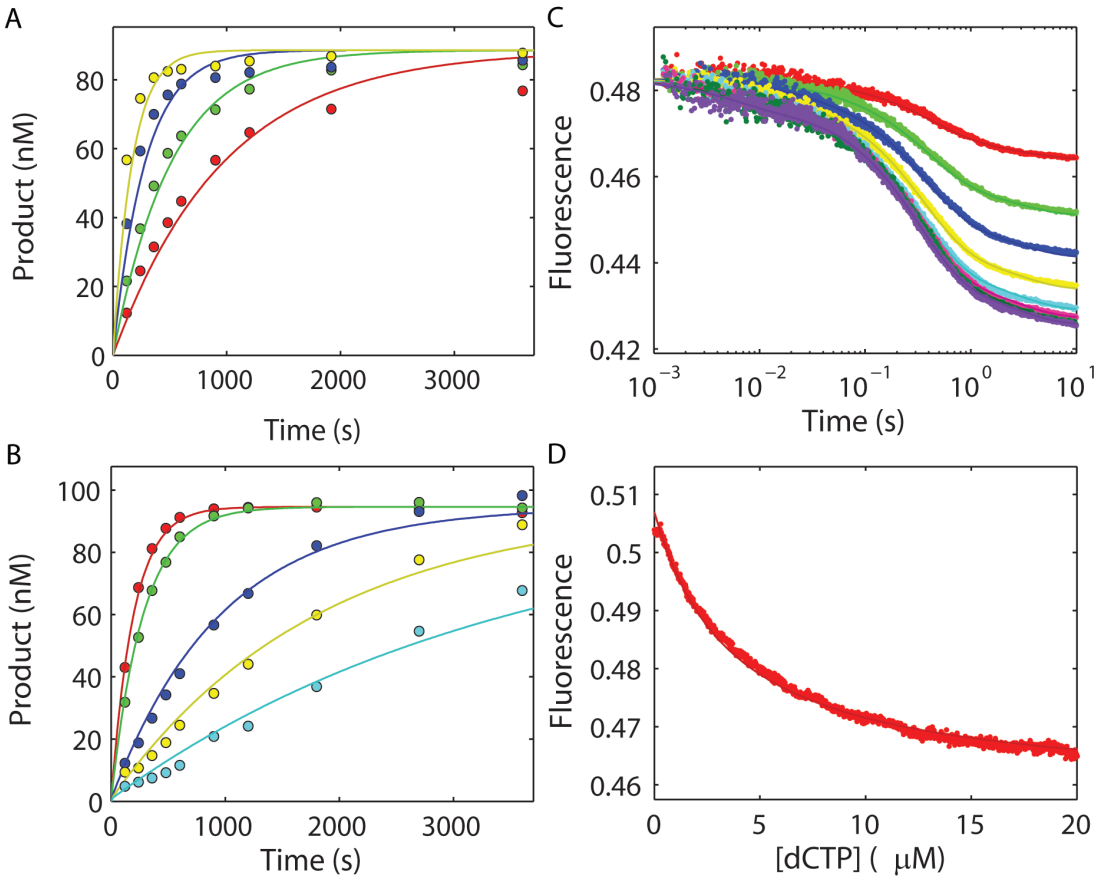



**Figure 3.8** *Global fitting of d4TMP excision and dCTP binding assays by MDCC-labeled HIVRT_TAMs_T69SSSVD.* ATP (A) and dCTP (B) concentration dependence of d4TMP excision and rescue as monitored by the formation of elongated product. (A) A preformed TAMs_T69SSSVD/26d4T complex (150 nM MDCC-labeled RT mutant and 100 nM of 26d4T/45mer) was rapidly mixed with 100 μM TTP and 1 μM dCTP in the presence of various concentrations of ATP (0.4, 0.8, 1.6, and 3.2 mM). (B) A preformed enzyme-DNA duplex (150 nM MDCC-labeled mutant and 100 nM of 26d4T/45mer) was rapidly mixed with 100 μM TTP and 4.8 mM ATP in the presence of various concentrations of dCTP (5, 10, 50, 100 and 200 μM). Binding affinity of dCTP to enzyme-DNA complex terminated with d4TMP was measured by ensemble stopped-flow (C), as well as equilibrium titration assays (D). (C) A preformed TAMs_T69SSSVD/26d4TMP complex (100 nM mutant and 150 nM of 26d4T/45mer) was rapidly mixed with various concentrations of dCTP (2, 5, 10, 20, 40, 60, 80 and 100 μM). Fluorescence was observed for 10 s (time was showed in logarithmic scale). (D) In the titration assay, the preformed TAMs_T69SSSVD/26d4TMP complex (100 nM mutant and 150 nM of 26d4T/45mer) was continuously mixed with increasing concentrations of dCTP (0-20 μM) while recording fluorescence. Smooth lines represent the best fit to get the rate constants shown in Table 3.2 and 3.3.

**Fingers Insertion Broadened Excision Specificity by Alleviating Inhibition from the Next Correct Nucleotide**

Previous studies have shown that fingers insertion enhanced ATP dependent excision of NRTIs in the presence of one or more TAMs mutations (*117, 124*). However, our data suggested that introduction of a Ser-Ser dipeptide in the fingers domain decreased excision efficiencies of ddTMP and d4TMP by 2~4 fold if one only considers $k_{exc}/K_{m,ATP}$ (Table 3.2). However, the fingers insertion also made the binding of dCTP weaker by 7~10 fold (Table 3.3), which therefore results in less inhibition against the excision of thymidine analogs (Figure 3.9 and Table 3.2). As shown in Table 3.2, the mutant containing a dipeptide insertion showed higher excision rates in the presence of 50 μM of dCTP, which is close to its physiological concentration of approximately 30 μM (*131*). To demonstrate this trend more clearly, excision rates of ddTMP and d4TMP by TAMs and TAMs_T69SSSVD mutants were plotted versus dCTP concentrations shown in logarithmic scale (Figure 3.9). Excision rates of the TAMs_T69SSSVD mutant decreased less significantly with increasing concentrations of dCTP, either in the case of ddTMP (Figure 3.9A) or d4TMP (Figure 3.9B) excision. The data were fit to a hyperbola function (Equation 3.2), and apparent $K_d$ values obtained were 8.6 ± 0.6 (for TAMs/ddT), 50 ± 4 (for TAMs_T69SSSVD/ddT), 1.1 ± 0.3 (for TAMs/d4T), and 5 ± 0.2 μM (for TAMs_T69SSSVD/d4T), as summarized in Table 3.4.

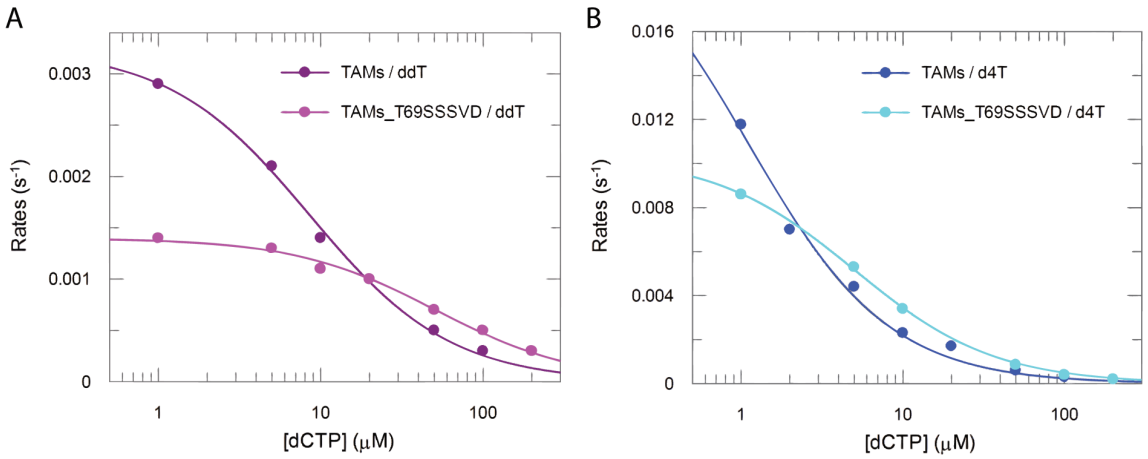



**Figure 3.9** *Fingers insertion alleviates inhibitory effect of dCTP on the excision of both ddTMP (A) and d4TMP (B) by TAMs and TAMs_T69SSSVD mutants.* Excision rates of ddTMP (A) and d4TMP (B) by TAMs and TAMs_T69SSSVD mutants were plotted against dCTP concentration shown on a logarithmic scale. Smooth lines in each panel represent conventional fitting of the concentration dependence of inhibition by dCTP to a hyperbola (Equation 3.2), which defines the apparent binding affinities of dCTP (apparent $K_d$) to TAMs/26ddTMP, TAMs_T69SSSVD/26ddTMP, TAMs/26d4TMP, and TAMs_T69SSSVD /26d4TMP complexes as 8.6 ± 0.6, 50 ± 4, 1.1 ± 0.3, and 5 ± 0.2 μM, respectively. The apparent weaker binding of dCTP onto TAMs_T69SSSVD/26ddTMP or 26d4TMP duplexes results in less inhibition at the physiological concentration of dCTP (~ 30 μM).

## 3.4 DISCUSSION

We have studied the effects of T69SSS fingers insertion of HIVRT on ATP mediated excision of thymidine analogs and the binding of next correct nucleotide (dCTP) onto the 3' end of thymidine analog-terminated primers. Previous studies have shown that fingers insertion in the backbone of TAMs could broaden the excision specificity by destabilizing the formation of dead-end complex (*115, 117, 120*). However, accurate measurements and comparisons of the excision efficiencies ($k_{exc}/K_{m,ATP}$) by different mutants are still missing since none of these studies explored the ATP concentration dependence of the reaction. Also, the binding affinities of next correct nucleotides were usually determined by gel mobility shift assays, which unfortunately, is not a reliable method due to the rapid dissociation of dCTP from the enzyme. We now revisit this problem and carefully determine the excision efficiencies, as well as apparent binding affinities of dCTP by measuring the concentration dependence of the reactions followed by robust global data analysis. Our data show that a Ser-Ser insertion decreased excision efficiencies of all thymidine analogs tested by 2-4 fold while weakened the binding of dCTP by 7-10 fold, which resulted in the enhanced excision ability by "69 insertion complex" observed under physiological conditions by a factor of ~1.5.

Fitting data based upon numerical integration of the rate equations with no simplifying assumptions is far more robust and demanding than fitting to equations derived from a simplified form of the model because it requires that the model account for both the rate and amplitude of all reactions simultaneously (*77, 86*). To address this, our data were analyzed by both conventional (Figure 3.2) and global fitting methods

(Figure 3.3-3.4), as detailed in *Materials and Methods*. Figure 3.2A & C show the ATP or dCTP concentration dependent rescue and extension assays, in which a purified ddTMP terminated primer was used. The excised chain terminators can interfere with the excision and extension reactions since the dinucleoside tetraphosphate product is a good substrate for the reaction to restore the analog (*142-143*). However, a purified substrate minimized the interference when the drug was removed due to its relatively low concentration (100 nM) yielding a small concentration of the dinucleoside tetraphosphate compared to that of TTP (100 μM), so re-formation of terminated primer was negligible. Also, the excision is the rate-limiting step, followed by a very fast extension of 2 nucleotides, which drives the net reaction to completion. The design of the assays ensures that we can obtain reliable data for further analysis. Hyperbolic analysis (Figure 3.2B and D) yields key parameters, such as $k_{exc}$, $K_{m,ATP}$, excision efficiency ($k_{exc}/K_{m,ATP}$), and the apparent binding affinity of dCTP, which can be used as first estimations in global analysis and to double check values obtained from global fitting. The two stopped flow assays (Figure 3.2E and G) provided more information about dCTP binding from other aspects. The amplitude changes of dCTP induced fluorescent signals are proportional to the fraction of bound dCTP, from which apparent binding affinity of dCTP could be obtained (Figure 3.2F). Surprisingly, the apparent $K_d$ values obtained from stopped flow assays are not consistent with the value obtained from quench flow assays which is ~5 fold tighter. This trend is more obvious when we globally fit the data set to a three-step binding model (Scheme 3.1), as shown in Figure 3.3. The scheme was derived based on the Boyer model (*79*), where the competition between drug removal and DEC formation

was governed by the equilibrium constant ($K_0$) of the reverse translocation between P site and N site. The failure of global fitting of the data inspired us to re-evaluate our model and ask if we miss certain important factors. The first possibility we examined is that maybe the inhibition by dCTP is not specific and the apparent tighter binding of dCTP is due to additional inhibition from other nucleotides in the reaction. This possibility was ruled out quickly since our data clearly showed that only dCTP could inhibit the reaction (data not shown), which is also consistent with results published by other groups (*127*). Then we carefully examined the differences between chemical quench and stopped flow assays and realized that there was no ATP in fluorescent measurements but the concentration was high in quench assays (4.8 mM). So a hypothesis was formulated that the apparent tighter binding of dCTP observed in quench flow assays is due to an additional conformational change induced by the binding of ATP to the E-$DNA_{dd}$-dCTP ternary complex. This hypothesis was supported by the fact that a decrease of fluorescence was observed with a rate of ~0.5 $s^{-1}$, when a pre-formed ternary complex was mixed with 4.8 mM of ATP (Figure 3.4B). As was expected, when mixed with buffer, an increase of fluorescence was observed (Figure 3.4A) since a new equilibrium was established after the dilution, suggesting that the binding of ATP induced an additional conformational change.

Based on these observations, a new model was derived, as shown in Scheme 3.2, in which ATP can bind to the $GD_{25}$ddT.dCTP triplex and lead to the apparent tighter binding of dCTP in the chemical quench assays. Theoretically, ATP can bind to the ternary complex at any stage of the binding (either ground state binding or

conformational changes). In this circumstance only binding to the $GD_{25}$ddT.dCTP triplex is allowed to propose the minimal model to account for the data when the global fitting is performed. As shown in Figure 3.5, this model matches the data set very well, from which parameters governing excision efficiencies and binding affinities of dCTP were determined, as summarized in Tables 3.2 and 3.3. For ddTMP excision by the TAMs mutant, the $k_{exc}$, $K_{m,ATP}$ and $k_{exc}/K_{m,ATP}$ values obtained from global fitting are consistent with corresponding values determined from conventional fitting (Figure 3.2B). This analysis shows how data global fitting provide more kinetic details and the inspiration as well as the means to solve contradictions observed in conventional data fitting.

Similar assays and data analysis were performed (Figure 3.6-3.8) to derive kinetic parameters governing excision efficiencies of thymidine analogs, as summarized in Table 3.2. A Ser-Ser insertion decreased the maximum excision rates and weakened the binding of ATP, which resulted in lower excision efficiencies. Usually, $K_m$ involves both binding and chemistry and therefore is not a pure parameter to describe binding affinity. Here $K_{m,ATP}$ was used as an estimation of the binding affinity of ATP since the excision is extremely slow (the highest rate is only 0.02 $s^{-1}$) so it is reasonable to presume that the binding of ATP reaches equilibrium before the chemistry occurs; thus, in this case the $K_{m,ATP}$ approximates the true $K_d$ of ATP (*77, 86-87*).

Extensive efforts were expended to derive the binding affinity of dCTP since the derivation is really a challenge in such a complex model. First, the translocation step and influence of ATP was put aside to derive the net dissociation constant of dCTP ($K_{d,dCTP} = 1/(K_1 \times (1 + K_2 + K_2K_3))$). As mentioned above, the high concentration of ATP makes the

binding of dCTP tighter in the quench flow assays since it induces an additional isomerization (governed by $K_{d,ATP}$). Then the apparent dissociation constant of dCTP in the presence of ATP can be determined as $K_{dCTP,app} = K_{d,dCTP} / (1 + [ATP] / K_{d,ATP})$. Finally, all the steps are taken into consideration since the equilibrium of the reverse translocation from the P site to the N site and the concomitant binding of ATP affect the fraction of E-$DNA_{dd}$ duplexes residing in the P site that are capable of the binding of dCTP. So the overall apparent dissociation constant of dCTP is $K_{d,app} = K_{dCTP,app} \times (1 + K_0 \times (1 + [ATP] /K_{m,ATP}))$. Here an example of how to derive the dissociation constant from a complex model was shown step by step and all the parameters were summarized in Table 3.3. Details of the derivation can be found in the *Materials and Methods*. But the formula can be much simpler in the case of the stopped flow assay since no ATP was added ($K_{d,app} = K_{d,dCTP} \times (1+K_0)$). The overall $K_{d,app}$ values for the binding of dCTP to the TAMs/26ddT duplex derived by this method are 1.3 ± 0.3 μM in the presence of ATP and 41 ± 10 μM in the absence of ATP, respectively. The former is 6-fold smaller than the number obtained from conventional fitting while the latter agrees very well with values determined from conventional analysis (with ATP, 8.6 ± 0.6 μM; without ATP, 41 ± 2 or 35 ± 0.1 μM [Figure 3.2 and Table 3.4]). Further analysis showed that, in the absence of ATP, the overall $K_d$,app values derived from global fitting are consistent with corresponding numbers obtained from conventional fitting of the stopped flow data, justifying the effectiveness of global fitting. In contrast, $K_d$,app values derived from global fitting are 2-6 fold smaller than the numbers obtained from conventional fitting of chemical quench data (Table 3.4).

**Table 3.4** *$K_{d,app}$ values obtained by conventional and global fitting*

| | $K_{d,app}$ With ATP (μM) | | $K_{d,app}$ Without ATP (μM) | | |
|---|---|---|---|---|---|
| | CF[a] | GF[b] | CF (SF[c]) | CF (ET[d]) | GF[b] |
| TAMs / 26ddT | 8.6±0.6 | 1.3±0.3 | 41±2 | 35±0.1 | 41±10 |
| TAMs_T69SSSVD / 26ddT | 50±4 | 9.9±2.5 | 80±6 | 50±0.2 | 90±20 |
| TAMs / 26d4T | 1.1±0.3 | 0.26±0.04 | 3.4±0.3 | 2.7±0.1 | 2.3±0.3 |
| TAMs_T69SSSVD / 26d4T | 5±0.2 | 2.5±0.6 | 3.9±0.1 | 4.6±0.2 | 4.6±0.8 |

[a] Conventional Fitting ; [b] Global Fitting; [c] Stopped Flow; [d] Equilibrium Titration.

This inconsistency of $K_{d,app}$ values demonstrated a potential drawback of the "$IC_{50}$ analysis" which has been widely used to determine apparent binding affinities of small inhibitors. This issue is more prominent in our cases first because dCTP is the substrate and inhibitor at the same time. More important is that the $K_d$ values to be determined are small numbers so it is almost impossible to achieve a dCTP concentration that is low enough to accurately determine $K_{d,app}$ while maintain high extension efficiency. For example, another group reported $IC_{50}$ values of 0.5±0.07 and 2.6±0.2 μM governing the inhibition against the rescue of d4TMP or ddTMP terminated primers (*144*). These numbers are 2-fold larger than what we obtained mainly because there are not enough reliable data points at low concentrations. Another issue of this research is that they separated excision from extension while ignoring the fact that the excision product is a good substrate for re-termination. In this situation, parameters obtained from global analysis are more reliable since the information contained in the data set defines a model, from which unknown information can be derived.

Our data show that the Ser-Ser insertion weakens the binding of dCTP by 2-fold in the absence of ATP. In quench assays, the dipeptide insertion can reduce the dCTP binding affinities by 7-10 fold, which leads to less extent of inhibition and therefore higher excision rates by the "69 insertion complex" under physiological conditions (Table 3.2 and Figure 3.9). Curves in Figure 3.9 demonstrate how lower excision efficiencies finally yield higher excision rates around the physiological concentration of dCTP. This trend is more obvious in the case of ddTMP excision by the TAMs_T69SSSVD because the apparent $K_d$ (9.9 ± 2.5 μM) is closer to the physiological concentration of dCTP (~30 μM).

The reverse translocation between P site and N site is a crucial step in our excision model, which determines the fate of a blocked primer – either DEC formation or rescue and extension. Site-specific foot-printing technique has been developed to study the translocation status of HIVRT in some studies to determine positions of cleavages (*72, 129-130*). However, the foot-printing reaction usually requires 2-5 minutes, which is too slow to accurately measure the rapid translocational equilibrium that can establish in less that 0.4 s (*20*). This method is not so convincing also because it was found that different reducing reagents (e.g., ascorbate or DTT) yielded different results (data not shown). On the other hand, the stopped flow data in this chapter provided sufficient information to define the equilibrium constant ($K_0$), as shown in Figure 3.10. We noticed that the fraction of fast conformational change was smaller when dCTP binds to the TAMs-E.$DNA_{d4T}$ duplex (Figure 3.10, red data points) than to the WT-E.$DNA_{d4T}$ complex (Figure 3.10, red solid line). This is because the TAMs mutations make RT cause the

DNA to shift more toward the N site so only a small fraction of protein-DNA complexes reside at the P site where nucleotides can bind directly. In our studies, we observed a slower reaction phase that represents the translocation from the N site to the P site. Therefore the ratio of the fraction of fast conformational change to that of the slower isomerization that is limited by the translocation step defines the equilibrium constant $K_0$. Previous single molecule studies performed in this lab showed that ~63% of ddTMP or d4TMP terminated primers locate at the P site of WT proteins, yielding a $K_0$ value of 0.6 (unpublished data). As summarized in Table 3.3, $K_0$ values obtained from these mutants are ranging from 3 to 12, suggesting that the TAMs mutations increased the fraction of DNA residing at the N site by 5-20 fold. Whether TAMs mutations enhance the excision of thymidine analogs by increasing the binding affinity of ATP or not is still controversial. However, the results presented here supported a new mechanism by which the TAMs mutations enhance the excision. Interestingly, an additional conformational change was also observed with extremely low rates (~0.05 $s^{-1}$) and very small amplitudes, suggesting that the binding of nucleotides to a analog-blocked primer can be different from their binding to a natural primer (*86*).

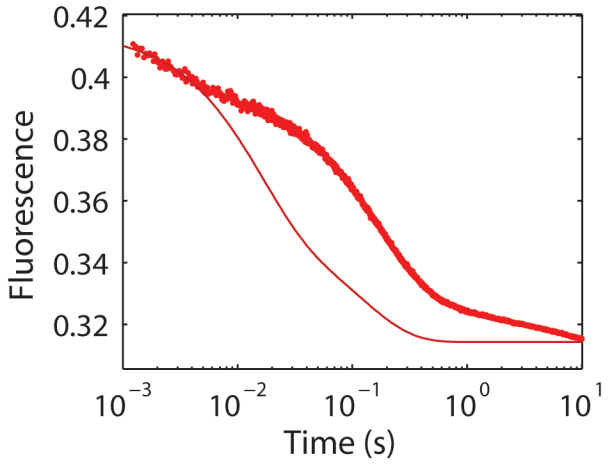



**Figure 3.10** *Representative data demonstrating the amplitude of the fast phase of fluorescence change defines the equilibrium of translocation between P site (Post-translocation) and N site (Pre-translocation).* The data were obtained by rapidly mixing a preformed TAMs/26d4TMP complex (100 nM TAMs and 150 nM of 26d4TMP/45mer) with 100 μM dCTP (green data points in Figure 3.7C) and shown in semi-logarithmic scale. The smooth line represents the fluorescence change upon mixing WT/26d4TMP duplex (100 nM WT RT and 150 nM of 26d4TMP/45mer) with 100 μM dCTP, which was generated with KinTek Explorer using previously published rate constants. The rate constants were kept unchanged but the scaling factors were adjusted to overlay the net reaction amplitude of the simulated curve with the real data. The fraction of the signal for the fast reaction phase provides an estimate of the equilibrium constant for the translocation reaction.

It has been proposed that the relatively high excision efficiency of AZT by the TAMs mutant is due to a fact that the long azido group can interfere with appropriate binding of the incoming nucleotide, so blocked primers have better access to the N site (*79*). However, another study indicated that the 3'-azido group alone could not confer high level of resistance and therefore is not the primary determinant responsible for its clinically observed high level of resistance (*80*). Our data showed that the TAMs mutant excises thymidine analogs with comparable efficiencies, 1.95 ± 0.05 $M^{-1}s^{-1}$ for AZT, 1.8 ± 0.1 $M^{-1}s^{-1}$ for ddT, and 8.1 ± 0.6 $M^{-1}s^{-1}$ for d4T, which are at least 10-fold higher than that of ddG (0.17 ± 0.06 $M^{-1}s^{-1}$, data not shown). The results clearly demonstrate that base components of nucleotide analogs determine the "original" excision efficiencies in the absence of inhibition. It seems that the TAMs mutations have evolved for the excision of thymidine analogs, which was supported by the fact that 3'-azido-guanosine (AZG) led to a different set of mutations during *in vitro* selection experiments, including L74V, F77L, L214F, K476N and V518I (*145-146*). On the other hand, azido group determines the extent to which the excision is resistant to the inhibition from the next correct nucleotide, which in turn determines the excision rates under physiological conditions. Our previous single molecule studies showed that only 14% of AZTMP terminated primers reside at P site, from which a $K_0$ value of 6 was derived (unpublished data). So the azido group shifts the equilibrium of translocation to favor the residence of an AZT-terminated primer at the N site by 10-fold. The TAMs mutations also facilitate a primer/template to locate at N site by 5-20 fold, varying with different chain terminators.

Consequently, the combination of these two factors resulted in the high level of excision of AZT by the TAMs mutant.

In this chapter, our data show that a Ser-Ser dipeptide insertion in the TAMs background decreased excision efficiencies of thymidine analogs and weakened the binding of next correct nucleotide. This weaker binding led to less extent of inhibition and therefore higher excision capacity at physiological concentrations of nucleotides. Our data also supported that the TAMs mutations facilitate the 3' end of a blocked primer binding in the N site, which is a novel mechanism by which these mutations result in enhanced excision. It was also found that the 3'-azido group was not the primary determinant for TAMs mediated excision and base components also played an important role to determine the excision specificity. Questions remain regarding the mechanistic basis for the >10-fold higher excision efficiencies of thymidine analogs by the TAMs mutant. An advanced understanding of how drug resistance mutations determine the excision selectivity could be used in the design of the next generation of NRTIs to overcome the evolution of resistance.

# Chapter 4: The Polymerase and RNase H Active Sites of HIV-1 Reverse Transcriptase Work Simultaneously and Independently

## 4.1 INTRODUCTION

HIVRT is a heterodimer, consisting of two related subunits: the catalytic p66 subunit, and a smaller p51 subunit that is proposed to play a structural role. The p66 subunit has two active sites, namely the N-terminal polymerase active site and the C-terminal RNase H active site, which perform different enzymatic functions: DNA-/RNA-dependent polymerization and RNase H cleavage, respectively. The two active sites are ~60 Å away from each other, flanking the nucleic acid binding cleft that spans 18 base pairs of a DNA/RNA hybrid (*15*). Despite the long distance, the process of reverse transcription requires a close coordination between the two catalytic centers, e.g., during minus-strand DNA synthesis, the polymerase active site copies the viral genome, while the RNase H site degrades the copied RNA segments to free the nascent DNA strand that serves as the template for plus-strand DNA synthesis. The fact that three activities of HIVRT evolved to be present in a single protein and that their coupled activities are required during reverse transcription suggested a possible spatial and temporal coordination between the two active sites. This hypothesis has been supported by several lines of evidence. First, RNase H is able to function as an independent enzyme, however, the isolated RNase H domain of HIV-1 RT was inactive, although its nuclease activity was able to be partially restored by adding various N- or C-terminal extensions (*147-152*), indicating other parts of RT were also required for the substrate binding and/or activity of RNase H. Secondly, the first RNA cleavage event occurs 18 base pairs from the 3'-OH

and seven-fold slower than DNA synthesis (*20*), suggesting a possible coupling between the two active sites. Subsequent RNase H cleavage events result in shorter products, but in the presence of a trap to inhibit substrate rebinding, only the 18-19 base products were observed, suggesting that sequential RNase H cleavage events are slow relative to dissociation of the DNA/RNA duplex from the enzyme (*92*). Finally, a more recent biochemical study suggested that HIVRT can simultaneously engage its DNA/RNA substrate at both active sites based on the observations that small ligands capable of stabilizing the nucleic acid in the polymerase active site only affected the pattern but not the efficiency of RNA cleavage under single-turnover conditions (*93*).

Despite the abundance of HIV-1 RT structures, only three of them were in complex with a DNA/RNA hybrid and none of them were in an RNase H competent mode (*15, 51, 54*). Therefore, the structural basis for the mechanism by which the two catalytic centers might coordinate with each other is still elusive. One study suggested that by positioning the DNA primer strand the 'RNase H primer grip' might also help to determine the trajectory of the hybrid substrate around the RNase H active site and thereby play an important role in coupling the two active sites (*15*). A later biochemical study then revealed that the cleavage specificity by RNase H could be modified by altering the RNase H primer grip (*50*).

Whether the two active sites of HIV-1 RT are able to act simultaneously is still under debate since structural and biochemical data suggested that primer extension and template degradation may be mutually exclusive and might occur sequentially. The DNA/RNA hybrids cocrystallized with HIV-1 RT were all in the polymerization mode

and did not reach the RNase H site (*15, 51, 54*). Meanwhile, the attempt to model a DNA/RNA hybrid into the RNase H active site of HIV-1 RT revealed that even with a severe bending, the substrate was not able to simultaneously occupy the two catalytic centers (*153*). These structural analyses suggested a 'toggling' model, where the bound substrate must toggle between the two active sites of HIV-1 RT, probably through conformational changes. As for the biochemical data, the rate of primer extension was determined to be 7-fold faster than that of RNA cleavage monitored simultaneously on a millisecond time scale (*20*) and the template was not degraded processively in RNA dependent extension assays (*19*), all of which were inconsistent with the tight temporal coupling between the two active sites. Moreover, RNA cleavage tended to occur more efficiently when primer extension paused before a hairpin on the RNA template (*94*). NNRTIs inhibitory to DNA synthesis were able to activate RNA cleavage, presumably through increasing the accessibility of the DNA/RNA hybrid substrate to the RNase H site (*95*). Based on these observations, it has been proposed that the nucleic acid substrate could only engage one active site at one time, but the data are not compelling.

In the present chapter, whether a DNA/RNA hybrid can simultaneously engage both active sites was further explored by examining single nucleotide incorporation and RNase H cleavage using pre-steady-state kinetics together with global data analysis. The data quantitatively demonstrated the existence of a spatial and temporal coordination between the two active sites. RNase H activities were also measured under the processive nucleotide incorporation condition and a model was proposed suggesting that the

polymerization activity of HIVRT was limited by the slow PPi release and therefore was able to act simultaneously with the RNase H activity at comparable rates.

## 4.2 MATERIALS AND METHODS

### Expression and Purification of Proteins Used in This Study

WT and mutant HIVRT proteins were expressed, purified and labeled following methods previously described (*77, 86*). Briefly, single mutations (W71A, W71D and R72A) were introduced in genes of both subunits using a Quikchange Multi kit (Stratagene). The method for the preparation of MDCC-labeled PBP used in this study was described previously (*154*) and its active-site concentration was determined by fluorescent titration with various concentrations of inorganic Pi.

### Nucleic Acid Substrates for Kinetic Studies

All DNA/RNA hybrids were designed to have a DNA primer recessed at the 3'-end on a RNA template to monitor the primary polymerase-dependent cleavage events (Table 4.1). The 45mer RNA templates with random sequence ($r45_{rd}$) or a string of Gs covering the RNase H cleavage site (r45) were purchased from Integrated DNA Technologies (IDT) with RNase free HPLC purification. The 25mer DNA primers complementary with $r45_{rd}$ ($d25_{rd}$) or r45 (d25) were ordered from IDT and purified by gel extraction.

**Table 4.1** *DNA/RNA Hybrids for Kinetic Assays*

*$d25_{rd}/r45_{rd}$:*

$d25_{rd}$: 5′-GCCTCGCAGCCGTCCAACCAACTCA-3′

$r45_{rd}$: 3′-CGGAGCGUCGGCAGGUUGGUUGAGU**A**GCAGCUAGGUUACGGCAGG-5′

*d25/r45:*

d25: 5′-GCCTCCCCCCCCCCAACCAACTCA-3′

r45: 3′-CGGAGGGGGGGGGGUUGGUUGAGU**A**GCAGCUAGGUUACGGCAGG-5′

A series of primers with various lengths (Figure 4.8B) were designed to study the minimum number of ribonucleotides downstream of the cutting site that are required for efficient cleavage. Shortening the primer by one nucleotide will shift the cutting site by one nucleotide towards the 3'-end of the RNA template in the DNA 3'-end directed cleavage mode. DNA primers terminated with ddAMP or ddTMP were made through enzymatic synthesis with RT. Following the biosynthetic reactions, terminated 25 or 26mer primers (25ddA or 26ddT) were purified by gel extraction. Primers and templates used in quench flow assays were 5' -$^{32}$P - labeled with T4 polynucleotide kinase (New England Biolabs). Annealing of DNA primers and RNA templates was carried out by mixing the two oligonucleotides at a 1:1 molar ratio and incubating at 67°C for 8 min, followed by slow cooling to room temperature.

**RNA Cleavage Assays**

Time courses of RNA cleavage were measured by rapid quench flow techniques using a KinTek RQF-3 (KinTek Corp.). 5'-labeled RNA template was annealed to the DNA primer as described above and diluted to a final concentration of 75 nM, pre-incubated with 175 nM of HIVRT and then rapidly mixed with 10 mM of $Mg^{2+}$ to initiate the cleavage. Effects of small-molecule ligands on polymerase-dependent RNase H

cleavage were studied by performing RNA cleavage assays in the presence of PFA (1 mM) or the next cognate nucleotides (500 μM, with dideoxy-terminated primers). To explore the potential coordination between the two active sites, incorporation and RNase H activities were measured simultaneously by pre-incubating 175 nM of HIVRT with 75 nM of double-labeled DNA/RNA hybrids and mixing with 100 μM TTP or ddTTP. All reactions were conducted in 50 mM Tris pH 7.5, 100 mM KCl (all concentrations final) at 37°C and were stopped by the addition of 0.5 M EDTA at various time points within 10 s. Products were separated by 15% denaturing PAGE and quantified using a Typhoon scanner in Imagine Quant 8.1 (GE). The incorporation and cleavage data were biphasic and fit with a double exponential equation to estimate the observed rates (conventional fitting).

RNase H activities were also explored under the processive nucleotide incorporation condition, in which all four nucleotides were added at 100 μM so the 25mer primer would be fully extended to the 45mer. In order to avoid the overlap of bands for extension and cleavage products, the two activities were assayed in parallel under the same conditions with the same enzyme by attaching the radioactivity only to one strand of the hybrid at one time, e.g., $^{32}$P was only labeled to the 5'-end of the primer when the incorporation activity was assayed and *vice versa*. Other procedures were the same except that data were collected for 20 s to allow full extension and multiple rounds of cleavage to occur.

**Global Data Analysis**

The inherent relationship between the rates and amplitudes of incorporation and cleavage reactions are crucial for the exploration of any potential coordination between the two active sites. However, this critical information can be lost in conventional fitting; therefore simultaneous fitting of all data within a given set is the preferred method. In the case of RNA cleavage during single nucleotide incorporation (TTP or ddTTP), the data were globally fit with two control experiments calibrating the RNase H activity of HIVRT against substrates before or after primer extension. The whole set of data were fit globally to a comprehensive mechanism showed in Scheme 4.1 using KinTek Explorer software (KinTek Corp.) (*155-156*). Three different RT-DNA/RNA complexes were identified and included in the model: $XD_{25}R_{45}$ (inactive), $YD_{25}R_{45}$ (polymerase active only), and $ED_{25}R_{45}$ (polymerase and RNase H both active), which underwent slow but reversible interconversions. Amplitudes of substrate depletion in primer extension (Figure 4.3A, red curve) and template degradation (Figure 4.3B, red curve) assays defined the equilibrium constants governing the exchanges between different species. Both $YD_{25}R_{45}$ and $ED_{25}R_{45}$ could extend the 25mer substrate to the 27mer product (with one mismatch extension), rates of which were derived from the nucleotide incorporation data (Figure 4.3A). RNA cleavage was able to occur before ($ED_{25}R_{45}$) or after ($ED_{26}R_{45}$) primer extension, yielding products with various lengths ranging from 35mer to 38mer. In globally fitting the data, cleavage rates were defined by the cleavage data obtained with (Figure 4.3B) or without (Figure 4.3C-D) the addition of an incoming nucleotide so that the cleavage pattern observed during polymerization was modeled based upon the

patterns seen with RNA/DNA duplex before and after elongation. In order to describe the complicated process as complete as possible, secondary cuts (e.g., $ED_{25}R_{38}$ to $ED_{25}R_{37}$) and nucleotide incorporation after template degradation (e.g., $ED_{25}R_{38}$ to $ED_{26}R_{38}$) were also included in the model. The simulation of RNA cleavage during ddTTP incorporation was similar except that the mismatch extension was prohibited and not included in the model (Scheme 4.2). The smooth curves in each figure panel represented the best simultaneous fit of all the data to the model and the rate constants were summarized in Scheme 4.1 (TTP) and 4.2 (ddTTP). Results of confidence contour analysis of these parameters were listed in Table 4.2.

Simulation of RNA cleavage during processive nucleotide incorporation was performed in an alternative way due to the complexity of reactions occurred in the process. First, the incorporation data were fit to a minimal model including sequential nucleotide incorporation events followed by possible product dissociation after each round of incorporation to account for the retention of a fraction of intermediates at long reaction times as attributed to their dissociation from the enzyme (Scheme 4.3 and Figure 4.9B). Global fitting of primer extension and RNA cleavage data was achieved by using rate constants obtained from the analysis of polymerization to constrain the simulation of RNA degradation data (Scheme 4.4 and Figure 4.9C). At the beginning, RT-DNA/RNA complexes were equilibrating between the RNase H active form ($ED_{25}R_{45}$) and inactive forms ($X(Y)D_{25}R_{45}$). Then, the $ED_{25}R_{45}$ complex underwent RNA cleavage (vertical arrows) before or after TTP incorporation (horizontal arrows) while the cleaved substrates were still extended. After the incorporation of multiple nucleotides, the second

round of RNA cleavage occurred, yielding shorter products. Then, another two rounds of RNA cleavage would occur until the primer was extended to full length and no more space was available for the polymerase-dependent RNA cleavage. In order to simplify the model, multiple incorporation events between different rounds of RNA cleavage were expressed as 'one' step kinetically, e.g., '+N8' represented the incorporation of 8 nucleotides in a row ($ED_{26}R_{45}$ to $ED_{34}R_{45}$). This simplification was appropriate and required because of the absence of intermediates. The cleavage products listed for each round (e.g., $ED_{25}R_{38}$ and $ED_{26}R_{38}$ for the first round) represented the sum of all cleavage products with various lengths and were expressed in that way for simplification. Rate constants derived from the simulation were summarized in Scheme 4.4 and Table 4.3.

## 4.3 RESULTS

### Kinetics of Nucleotide Incorporation and RNase H Cleavage.

In order to measure nucleotide incorporation and polymerase-dependent RNase H cleavage activities simultaneously, a double-labeled DNA/RNA substrate with a 25mer primer recessed at the 3'-end on a 45mer RNA template was designed (d25/r45, Table 4.1 and Figure 4.1A). Since sequence preferences have been observed for DNA 3'-end-directed cleavage (*157*), a string of Gs covering the cleavage window was placed in the template to avoid effects due to nucleotide preferences. Primer extension and template degradation were then studied simultaneously by rapid quench flow method under pre-steady-state conditions and both activities were observed to occur within the same time frame (Figure 4.1B). During TTP incorporation, RNA cleavage products mainly accumulated at positions -18 and -17, with minor bands formed at positions -16 and -15.

It was reported that the high-resolution site-specific footprinting was indicative of the translocational status of HIVRT(*129*), so we initially thought that the cleavage pattern observed here may have also reflected translocation stages of HIVRT on the DNA/RNA hybrid before and after nucleotide incorporation. However, control experiments performed in the absence of nucleotides revealed that the DNA 3'-end-directed RNA cleavage failed to strictly follow the translocation status of RT (Figure 4.2A). Biased cleavage to position -18 was observed for substrate d25/r45. When 26mer primers were used, cleavage products accumulated at 4 distinct positions instead of two that were characteristic of the pre- and post-translocation stages. Product formation was plotted versus time and both the incorporation and cleavage data were biphasic (Figure 4.1C and D). So rates of TTP incorporation and concomitant RNA degradation were determined by fitting the data to a double-exponential equation. The results showed that TTP was incorporated at a rate of $107 \pm 11\ s^{-1}$, which was 6~7-fold faster than the template degradation that was $17 \pm 2\ s^{-1}$. This difference was consistent with our previous data (*20*). Cleavage rates measured in the absence of nucleotides were determined as $7.7 \pm 1.7\ s^{-1}$ and $10.2 \pm 1.9\ s^{-1}$ for the 25mer and 26mer primers, which were a little bit slower than the cleavage during TTP incorporation (Figure 4.1D and 4.2B). Rates for the slower phase were around $0.5\ s^{-1}$, which were comparable in rate to DNA/RNA dissociation and rebinding to the enzyme.

Incorporation and cleavage activities were observed simultaneously but this observation did not exclude a sequential polymerization - hydrolysis model, because incorporation was faster and therefore could have been completed prior to the RNA

cleavage. In order to test this scenario, the rate of nucleotide incorporation was slowed by using a TTP analog (ddTTP) while exploring the RNA cleavage activity during ddTTP incorporation. A similar cleavage pattern was observed, with the major or minor products accumulated at positions -18 and -17, or -16 and -15, respectively (Figure 4.1B, right panel). The rate of ddTTP incorporation ($5 \pm 0.3\ s^{-1}$) was determined to be much slower than TTP incorporation ($107 \pm 11\ s^{-1}$), but was slightly slower than the RNA cleavage ($12 \pm 0.7\ s^{-1}$). Concomitant polymerase and RNase H activities were still observed even if they had comparable rates, suggesting that the two active sites were able to function simultaneously and independently. Moreover, in the single turnover experiment, if the two reactions occurred sequentially, one would expect to see a lag in the second reaction that is a function of the rate of the first reaction.

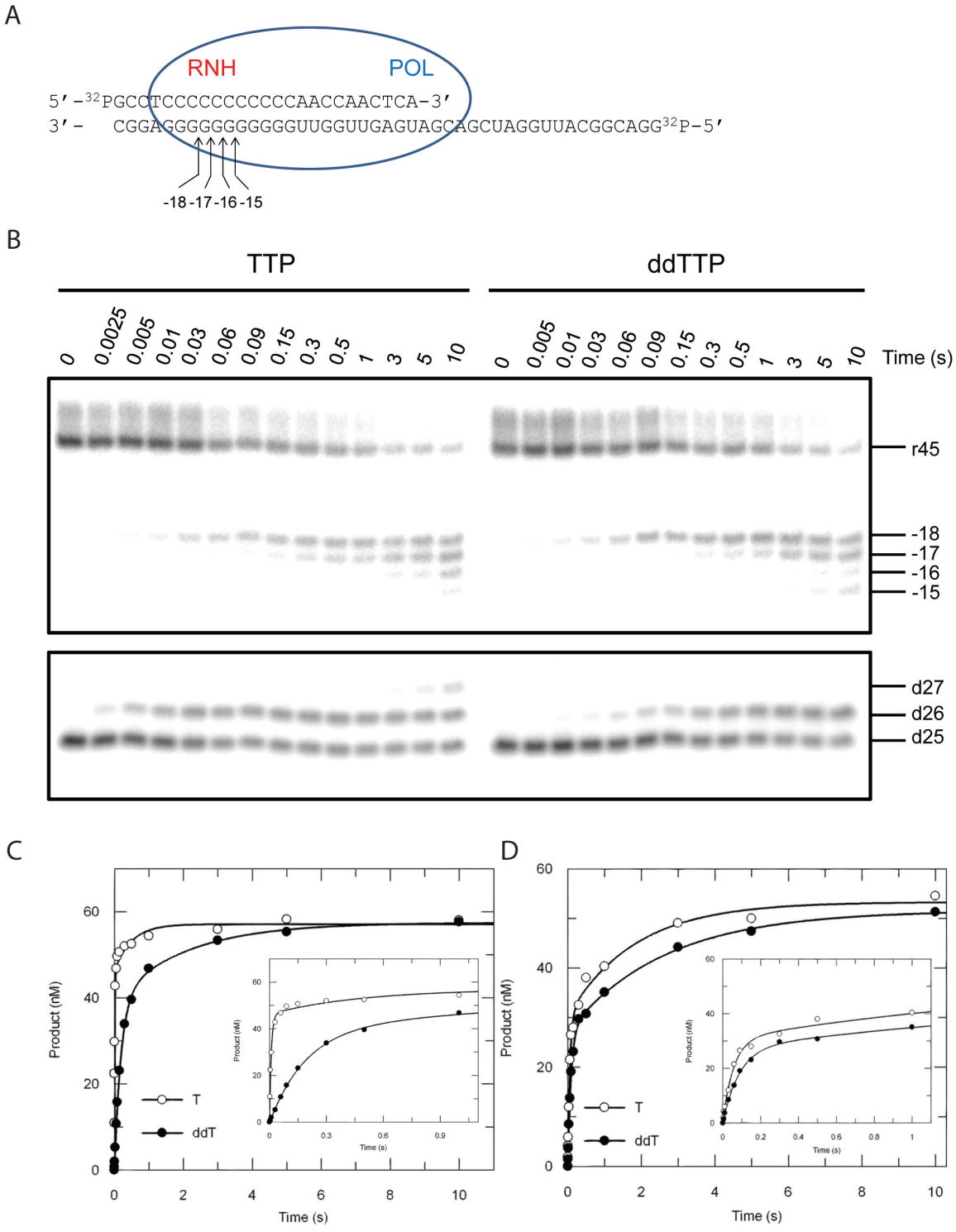

A
RNH
POL
5′-32PGCCTCCCCCCCCCCCAACCAACTCA-3′
3′- CGGAGGGGGGGGGGGGUUGGUUGAGUAGCAGCUAGGUUACGGCAGG32P-5′
-18 -17 -16 -15
B
TTP
ddTTP
0 0.0025 0.005 0.01 0.03 0.06 0.09 0.15 0.3 0.5 1 3 5 10
0 0.005 0.01 0.03 0.06 0.09 0.15 0.3 0.5 1 3 5 10
Time (s)
r45
-18
-17
-16
-15
d27
d26
d25
C
D
Product (nM)
Time (s)
T
ddT

**Figure 4.1**

**Figure 4.1** *Kinetics of Nucleotide Incorporation and RNase H Cleavage.* (A) The 'static' nomenclature for naming the cleavage sites are shown in the cartoon. The sequence of the double-labeled DNA/RNA hybrid is shown. The blue oval represents HIVRT, where POL and RNH denote the polymerase and RNase H active sites, respectively. Positions of cleavage sites are counting from the 3'-terminus of the DNA primer, e.g., -18 means the cleavage site is 18 nucleotides upstream of the 3'-end of the DNA primer. (B) To explore the potential coordination between the two active sites, polymerase (lower panel) and RNase H (upper panel) activities were measured simultaneously by pre-incubating 175 nM of HIVRT with 75 nM of double-labeled DNA/RNA hybrid and then mixing with 100 μM of TTP or ddTTP, in the presence of 10 mM $Mg^{2+}$. Reactions were stopped by adding 0.5 M EDTA at various time points up to 10 s. The 45mer RNA template and RNase H cleavages at positions -18 to -15 are marked. Positions of the 25mer primer and extended products are marked as well. Incorporation (C) and cleavage (D) data were analyzed by conventional fitting. The insets better display data points within 1 s. Nucleotide incorporation was fit to the double-exponential equation to derive rates of TTP or ddTTP incorporation as $107 \pm 11$ $s^{-1}$ and $5 \pm 0.3$ $s^{-1}$ for the fast phase, and $2 \pm 0.8$ $s^{-1}$ and $0.5 \pm 0.1$ $s^{-1}$ for the slower phase, respectively. Similarly, rates of RNA cleavage during TTP or ddTTP incorporation were determined to be $17 \pm 2$ $s^{-1}$ and $12 \pm 0.7$ $s^{-1}$ for the fast phase, and $0.5 \pm 0.1$ $s^{-1}$ and $0.4 \pm 0.04$ $s^{-1}$ for the slower phase.

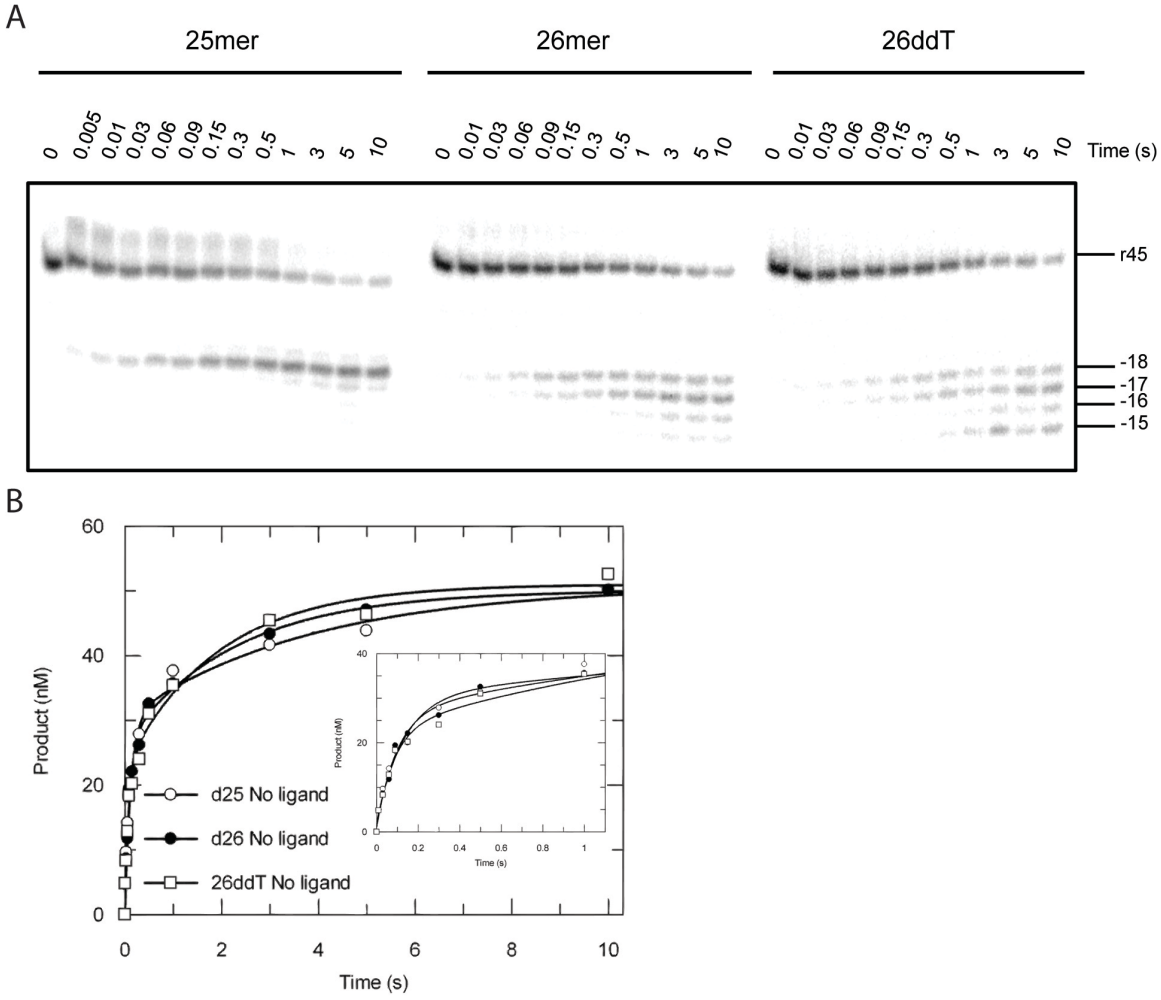



**Figure 4.2** *RNA Cleavage of Substrates before or after primer extension.* (A) The 45mer RNA template (r45) was annealed to 25mer, 26mer, or 26ddT primers to mimic substrates before or after nucleotides (TTP or ddTTP) incorporation. RNA cleavage assay of each substrate was performed as mentioned in Figure 4.3 and 4.4. The 45mer RNA template and RNase H cleavages at positions -18 to -15 are marked. (B) The Data of RNA cleavage against various DNA/RNA hybrids were analyzed by conventional fitting and the inset better display data points less than 1 s. Cleavage data were fit to the double-exponential equation to derive rates measured with d25/r45, d26/r45, and 26ddT/r45 as $7.7 \pm 1.7\ s^{-1}$, $10.2 \pm 1.9\ s^{-1}$, and $12 \pm 3\ s^{-1}$ for the fast phase, and $0.3 \pm 0.1\ s^{-1}$,$0.4 \pm 0.1\ s^{-1}$, and $0.5 \pm 0.1\ s^{-1}$ for the slower phase, respectively.

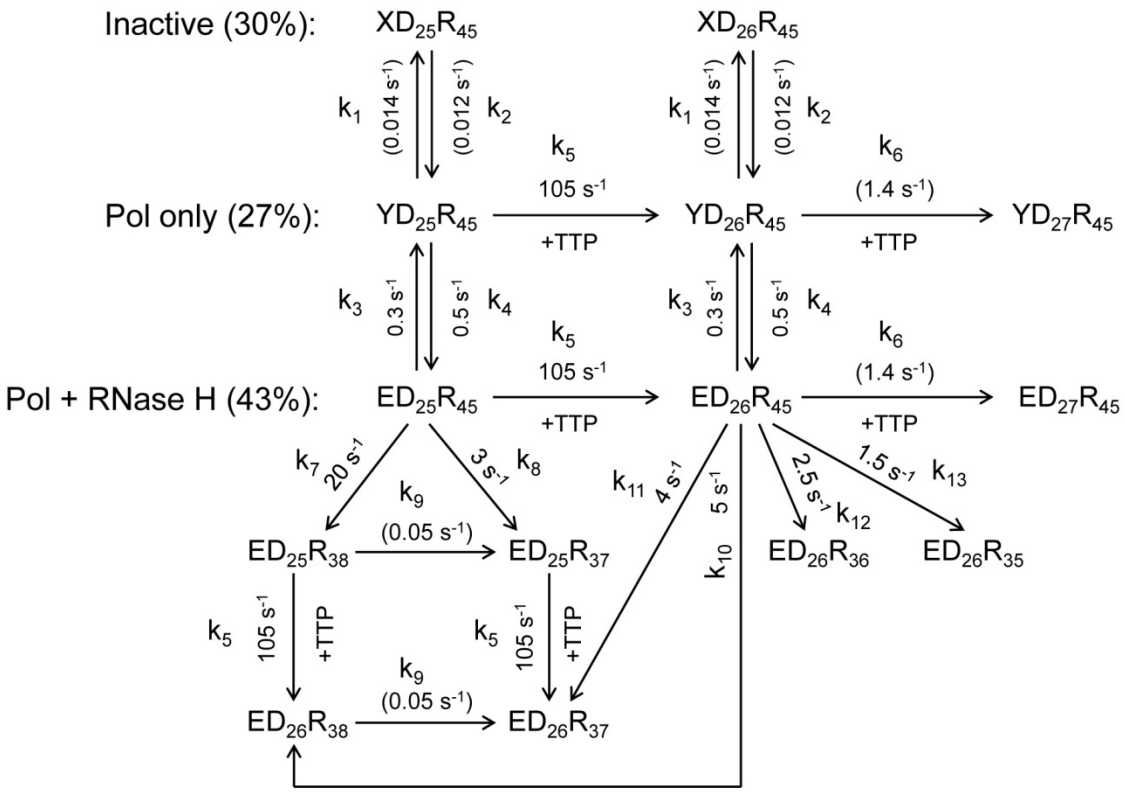
**Global Data Fitting of Nucleotide Incorporation and RNase H Cleavage**

Conventional fitting of above data provided some evidence that the two catalytic centers of HIVRT work independently. However, conventional analyses were unable to rule out the possibility that the simultaneous activities observed previously were due to mixed enzyme states where one enzyme catalyzes polymerization while another catalyzes RNase H hydrolysis. Information to distinguish whether a single enzyme species catalyzes both reactions simultaneously lies in analysis of the amplitudes of the reactions. The inherent relationships between the rates and amplitudes of the incorporation and cleavage reactions provide more information to determine with both reactions are occurring simultaneously and this can be accomplished using global data fitting. Global data analysis was first performed for the 4 experiments measuring TTP incorporation (Figure 4.1B, upper left), concomitant RNA cleavage (Figure 4.1B, bottom left), and RNase H activities of RT against substrates before or after (Figure 4.2A) primer extension.

Data were fit using a minimal model shown in Scheme 4.1, as described in *materials and methods*. Distinct from the conventional data fitting where the amount of products with various lengths were summed to define the net reaction rate, for global fitting each band was quantified and analyzed individually. Figure 4.3A shows the time course of primer extension during TTP incorporation, while Figure 4.3B shows the simultaneous accumulation of RNase H cleavage products. Figures 4.3C and D show the time course of RNase H cleavage before and after primer extension, respectively. The smooth lines represent the global fit to the entire data set according to Scheme 4.1.

The comprehensive mechanism showed in Scheme 4.1 was derived as a minimal model to account for the data. First, neither the 25mer primer (Figure 4.3A, red curve) nor the 45mer RNA template (Figure 4.3B, red curve) were decayed with a single fast rate, suggesting the existence of a nonproductive form of RT-DNA/RNA complexes ($XD_{25}R_{45}$). Second, the amplitude of the fast phase of 25mer consumption (from polymerization) was approximately the same as the sum of amplitudes of the fast and slower phases of 45mer depletion (RNase H), suggesting that a fraction of polymerase-competent RT-DNA/RNA complexes were not active for RNase H cleavage ($YD_{25}R_{45}$) and the slower phase of 45mer consumption was due to the delayed switch from the inactive form ($YD_{25}R_{45}$) to the active form ($ED_{25}R_{45}$). Our data also clearly supported the existence of the bifunctional form of RT-DNA/RNA complexes that could perform both activities simultaneously ($ED_{25}R_{45}$). Equilibrium between the three species were determined by the ratios of the amplitudes: the ratio of amplitudes of fast and slower phases of TTP incorporation determined the equilibrium between the polymerase active and inactive forms ($YD_{25}R_{45}$ + $ED_{25}R_{45}$, 70%; $XD_{25}R_{45}$, 30%), whereas the ratio of amplitudes of fast and slower phases of RNA cleavage determined the equilibrium between the bifunctional and polymerase-only forms ($ED_{25}R_{45}$, 43% ; $YD_{25}R_{45}$, 27%). Rates of conversions among those species were listed in Scheme 4.1 and Table 4.2. The results showed that conversions between the active and inactive forms were very slow, with the forward and reverse rates of ~0.014 and ~0.012 $s^{-1}$, respectively. The polymerase-only form ($YD_{25}R_{45}$) switched to the bifunctional form ($ED_{25}R_{45}$) at a rate of 0.5 $s^{-1}$, which was in good agreement with the rate of slower phase (0.5 ± 0.1 $s^{-1}$)

obtained from the conventional fitting of the cleavage data (Figure 4.1D). Incorporation rates were determined as 105 $s^{-1}$ for correct base pairing and 1.4 $s^{-1}$ for formation of a mismatch (27mer). Rates of RNA cleavage before primer extension were 20 $s^{-1}$ at position -18 and 3 $s^{-1}$ at position -17, consistent with the strong bias towards -18 cleavage observed in the control experiment (Figure 4.2). Rates of RNA cleavage after primer extension were 5, 4, 2.5, and 1.5 $s^{-1}$ for cleavages at positions -18 to -15, respectively. The sum of all rates in a branched pathway yielded rates of RNA cleavage before and after primer extension as 23 and 13 $s^{-1}$, respectively. Therefore, the observed rate of 17 ± 2 $s^{-1}$ determined from the conventional fitting (Figure 4.1D) appears to be a composite rate combining these two reactions. Secondary cuts also occurred since cleavage products corresponding to position -18 increased with time at the beginning but eventually decreased (Figure 4.3B and D, green curves), the rate of which, however, was much slower (~0.05 $s^{-1}$), suggesting that it was highly inefficient.

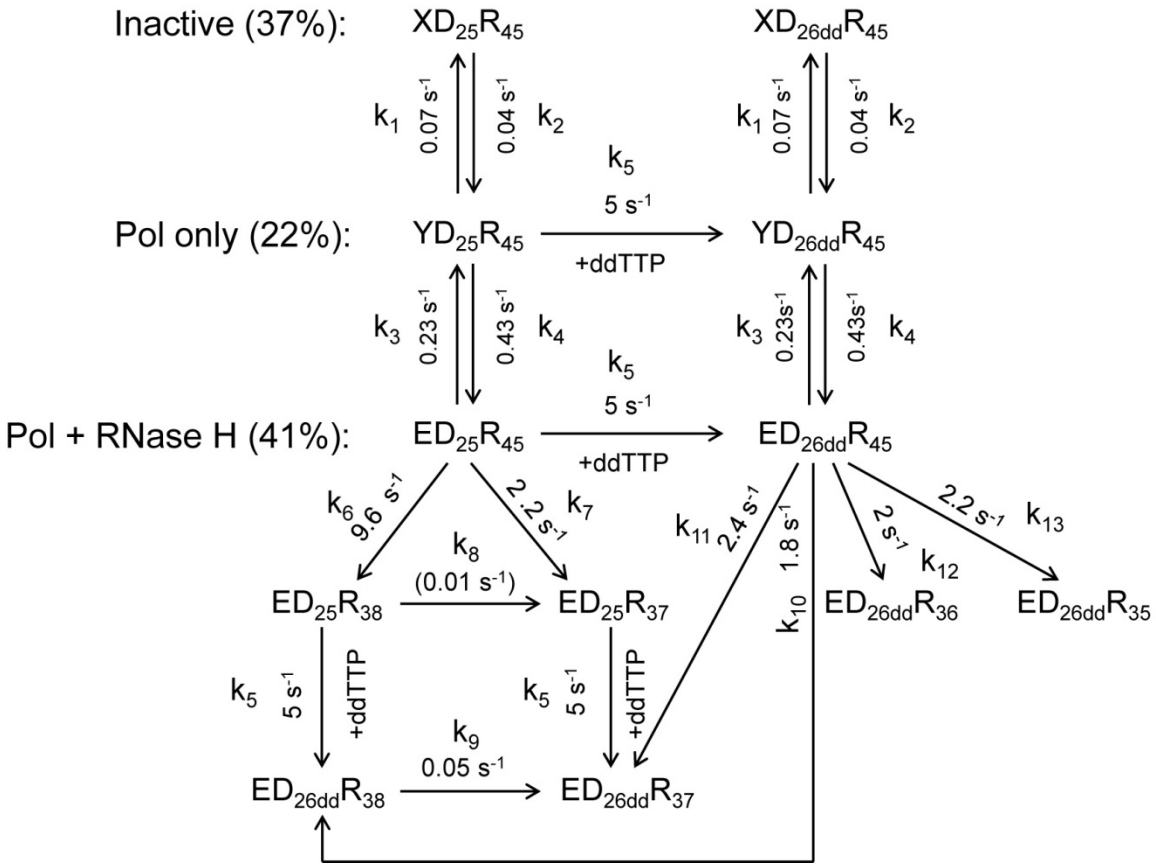



**Scheme 4.1** *Simultaneous TTP Incorporation and RNA Cleavage.*
A minimal model for simultaneous TTP incorporation and RNA cleavage is shown above. Three different RT-DNA/RNA complexes were identified and included in the model: $XD_{25}R_{45}$ (inactive), $YD_{25}R_{45}$ (polymerase active only), and $ED_{25}R_{45}$ (polymerase and RNase H both active). These three species underwent slow but reversible inter-conversion. Both $YD_{25}R_{45}$ and $ED_{25}R_{45}$ could extend the 25mer substrate to the 27mer product (including one mismatch incorporation). RNA cleavage was able to occur before ($ED_{25}R_{45}$) or after ($ED_{26}R_{45}$) primer extension, yielding products with various lengths ranging from 35mer to 38mer. In order to describe the complicated process as completely as possible, secondary cuts (e.g., $ED_{25}R_{38}$ to $ED_{25}R_{37}$) and nucleotide incorporation after template degradation (e.g., $ED_{25}R_{38}$ to $ED_{26}R_{38}$) were also included in the model. Rate constants for each step are shown in the model and labeled by $k_n$ (for n = 1 to 13).

Similar data analysis was performed for RNA cleavage during ddTTP incorporation (Figure 4.4), in order to explore the effect of a slower polymerization rate on the RNase H activity. Rate constants governing the equilibrium between various RT-DNA/RNA complexes, nucleotide incorporation, and RNA cleavage were derived from the global data fitting and summarized in Scheme 4.2 and Table 4.2. Similar to our analysis of TTP incorporation, the nonproductive (37%), polymerase-competent only (22%), and bifunctional species (41%) of RT-DNA/RNA complexes were also identified. The rate of ddTTP incorporation was determined as 5 $s^{-1}$, which was consistent with the number obtained from conventional fitting (5 ± 0.3 $s^{-1}$). The rate of RNA cleavage prior to ddTTP incorporation was 11.8 $s^{-1}$, equal to the sum of 9.6 and 2.2 $s^{-1}$ to form -18 and -17 cleavages, respectively. A net rate of 8.4 $s^{-1}$ was determined for RNA cleavage after primer extension, with comparable rates derived for cleavages at various positions (-18: 1.8 $s^{-1}$, -17: 2.4 $s^{-1}$, -16: 2 $s^{-1}$, and -15: 2.2 $s^{-1}$). The cleavage rates were also consistent with the number (12 ± 0.7 $s^{-1}$) derived from the conventional fitting (Figure 4.1D). Again, the secondary cuts occurred at very slow rates of ~0.01 and 0.05 $s^{-1}$, respectively.

Our global data analysis demonstrated that nucleotide incorporation and RNA cleavage occur simultaneously under single-turnover conditions, and we have defined the fraction of RT-DNA/RNA complexes that had the substrate engaged at both active sites. Also, RNA cleavage could take place at the same time as nucleotide incorporation even if rates of the two activities were comparable, which disproved the sequential model where one active center acted faster and therefore could have accomplished its tasks prior to the second catalytic center. Moreover, we show the coincident appearance of multiple

cleavage bands, providing evidence for some plasticity in the location of the cleavage site. For example during incorporation of either TTP or ddTTP, we see the nearly equal formation of -18 and -17 cleavage products, but see no evidence for the cleavage at position -18 before and at -17 after nucleotide incorporation (Figure 4.1B). Note that our nomenclature for naming the cleavage site is static (Figure 4.1A), so that after primer extension, cleavage 18 base-pairs from the new primer terminus results in a -17 product, numbered based upon the starting material. This nomenclature was chosen so that a given band is consistently named even when we do not know the primer length during a given cleavage event when measuring both reactions simultaneously (as in Figure 4.1).




**Scheme 4.2** *Simultaneous ddTTP Incorporation and RNA Cleavage.*
The minimal model for simultaneous ddTTP incorporation and RNA cleavage is shown above. Three different RT-DNA/RNA complexes were identified and included in the model: $XD_{25}R_{45}$ (inactive), $YD_{25}R_{45}$ (polymerase active only), and $ED_{25}R_{45}$ (polymerase and RNase H both active), which underwent slow but reversible inter-conversion. Both $YD_{25}R_{45}$ and $ED_{25}R_{45}$ could extend the 25mer substrate to the dideoxy-terminated 26mer product (mismatch extension was blocked). RNA cleavage was able to occur before ($ED_{25}R_{45}$) or after primer extension ($ED26_{dd}R_{45}$), yielding products with various lengths ranging from 35mer to 38mer. In order to describe the complicated process as completely as possible, secondary cuts (e.g., $ED_{25}R_{38}$ to $ED_{25}R_{37}$) and nucleotide incorporation after template degradation (e.g., $ED_{25}R_{38}$ to $ED26_{dd}R_{38}$) were also included in the model. Rate constants for each step are shown in the model and labeled by $k_n$ (for n = 1 to 13).

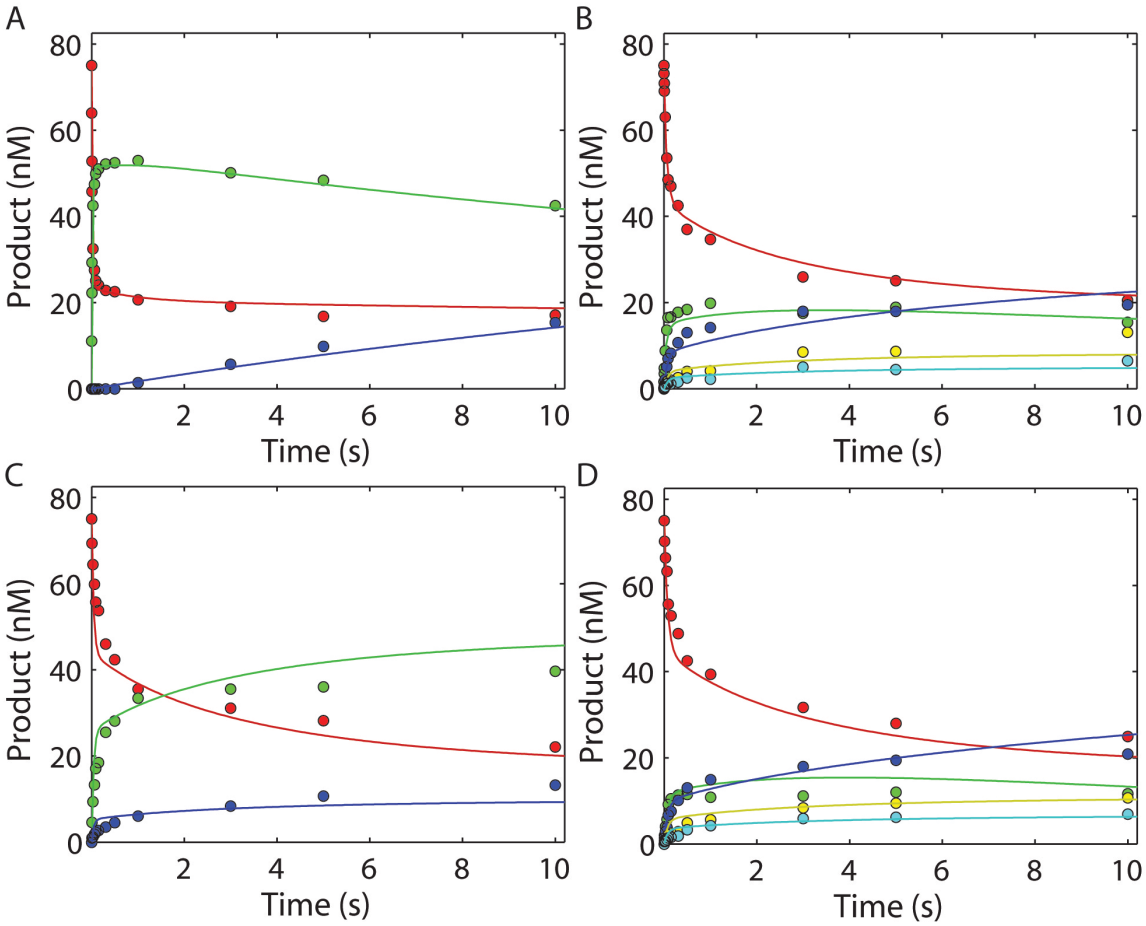



**Figure 4.3** *Global Data Fitting of TTP Incorporation and RNase H Cleavage*. Data for TTP incorporation (A) and concomitant RNase H cleavage (B) were analyzed by global data fitting. Two control experiments were performed to measure the RNase H activities of HIVRT against substrates before (C) or after (D) primer extension. Briefly, pre-formed RT-DNA/RNA complexes (175 nM of HIVRT and 75 nM of d25/r45 (C) or d26/r45 (D) ) were rapidly mixed with 10 mM $Mg^{2+}$, in the absence of any nucleotides. All the data were fit to a single model using KinTek Explorer to derive the rate constants summarized in Scheme 4.1 and Table 4.2. Smooth lines in each panel represent the best fit to the model. Each band was quantified individually so more information can be derived from the global fitting. Color codes: (A) 25mer primer, red; 26mer product, green; and 27mer product, blue. (B-D) 45mer RNA template, red; cleavage at position -18 (38mer), green; cleavage at position -17 (37mer), blue; cleavage at position -16 (36mer), yellow; and cleavage at position -15 (35mer), cyan.

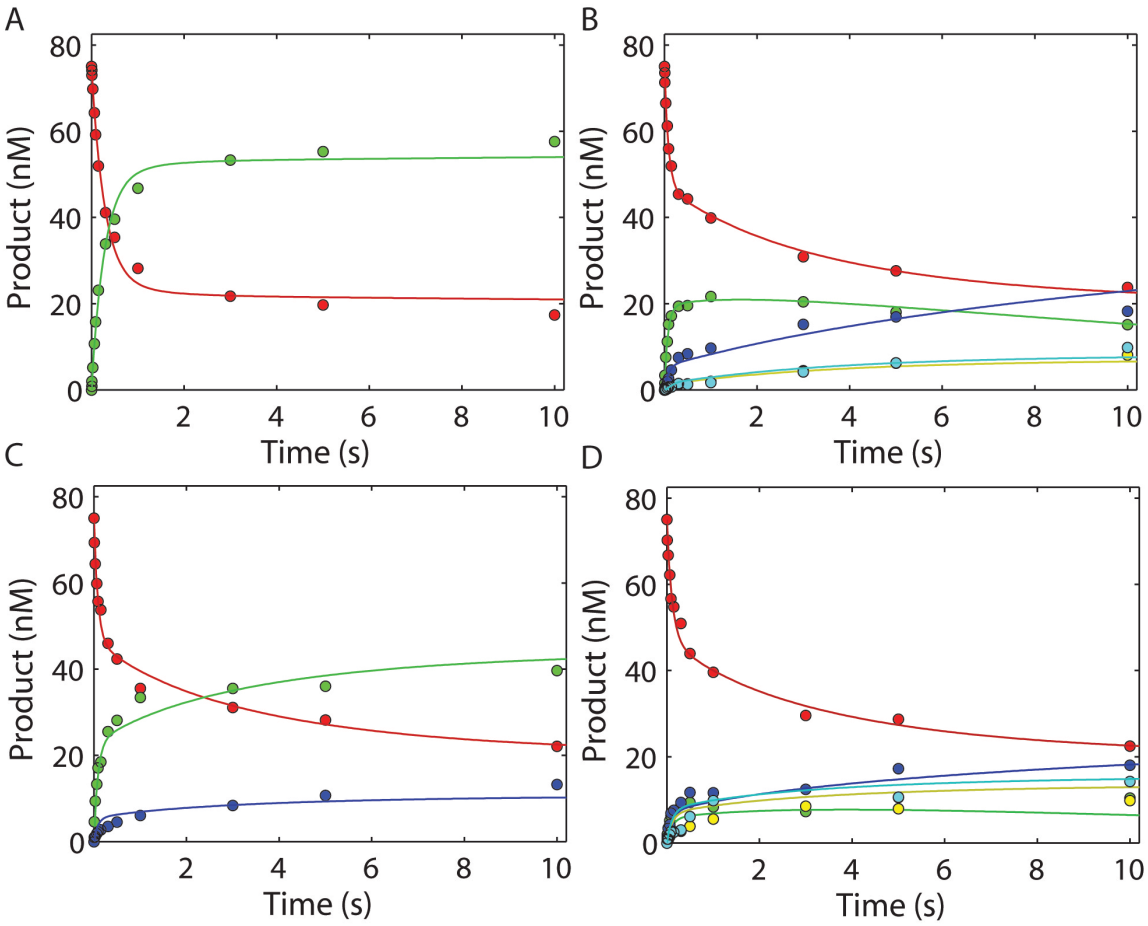



**Figure 4.4** *Global Data Fitting of ddTTP Incorporation and RNase H Cleavage*. Data of ddTTP incorporation (A) and concomitant RNase H cleavage (B) were analyzed by global data fitting. Two control experiments were performed to measure the RNase H activities of HIVRT against substrates before (C) or after (D) primer extension. Briefly, pre-formed RT-DNA/RNA complexes (175 nM of HIVRT and 75 nM of d25/r45 (C) or d26ddT/r45 (D) were rapidly mixed with 10 mM $Mg^{2+}$, in the absence of any nucleotides. All the data were fit to a single model using KinTek Explorer to derive rate constants summarized in Scheme 4.2 and Table 4.2. Smooth lines in each panel represent the best fit to the model. Each band was quantified individually so more information can be derived from the global fitting. The same color codes were used here as in Figure 4.3.

**Table 4.2** *Kinetic Parameters Governing Nucleotide Incorporation and RNase H Cleavage*[a]

| Nucleotides | $k_1$ ($s^{-1}$) | $k_2$ ($s^{-1}$) | $k_3$ ($s^{-1}$) | $k_4$ ($s^{-1}$) | $k_5$ ($s^{-1}$) | $k_6$ ($s^{-1}$) | $k_7$ ($s^{-1}$) |
|---|---|---|---|---|---|---|---|
| TTP | 0.014 (0.009~0.05)[b] | 0.012 (0.008~0.04)[b] | 0.3 (0.18~0.36) | 0.5 (0.30~0.63) | 105 (67~152) | 1.4 (0.4~2.5)[c] | 20 (8~24) |
| ddTTP | 0.07 (0.04~0.09) | 0.04 (0.02~0.05) | 0.23 (0.18~0.30) | 0.43 (0.33~0.56) | 5 (4.9~6.2) | 9.6 (8.6~10.9) | 2.2 (1.6~2.7) |
| | $k_8$ ($s^{-1}$) | $k_9$ ($s^{-1}$) | $k_{10}$ ($s^{-1}$) | $k_{11}$ ($s^{-1}$) | $k_{12}$ ($s^{-1}$) | $k_{13}$ ($s^{-1}$) | |
| TTP | 3 (1.3~4.9) | 0.05 (0.007~0.09)[c] | 5 (3.6~7.4) | 4 (3.4~5.6) | 2.5 (1.9~3.1) | 1.5 (0.9~2.2) | |
| ddTTP | 0.01 ($\leq$ 0.03)[c] | 0.05 (0.03~0.06) | 1.8 (0.9~2.8) | 2.4 (1.6~3.8) | 2 (1.3~3.1) | 2.2 (1.4~3.5) | |

[a] The lower and upper limits of each rate constant derived from confidence contour analysis of the simulation were included in the parenthesis. The nomenclature of rate constants was illustrated in Scheme 4.1 and 4.2.

[b] Individual rate constant was not well defined by the data, as illustrated by the relatively wide range. The ratio of the two (the equilibrium constant), however, was well constrained.

[c] Rate constants governing those slow reactions were not well defined by the data. Therefore, either a wide range or an upper limit is specified.

**Effects of Small Ligands on Polymerase-dependent RNase H Cleavage**

Previous structural analyses have suggested a 'toggling' model, where the bound substrate must toggle between two active sites of HIVRT, probably through conformational changes (*15, 51, 54, 153*). However, a recent biochemical study provided data to argue against this mechanism based on observations that small molecule ligands capable of stabilizing the nucleic acid in the polymerase active site only affected the pattern but not the efficiency of RNA cleavage under single-turnover conditions (*93*). This issue was revisited here since a quantitative analysis of the previously published cleavage data was missing. In addition, a study performed with the substrate without sequence bias can be more definitive.

The pyrophosphate analogue PFA and next correct nucleotide (with dideoxy-terminated primers) were previously reported to be able to stabilize the pre- or post-translocated RT-DNA complexes (*129, 158*) and presumably could lock the nucleic acid substrate in the DNA polymerase active site. Effects of these ligands on polymerase-dependent RNase H activities were explored with various DNA/RNA hybrids containing 25mer or 26mer primers, in the absence or presence of ligands (Figure 4.5A and B). When 25mer primers were examined, predominant cleavage products accumulated at position -18 in the absence of a ligand, whereas PFA shifted more than half of the cleavages towards position -19, and saturating concentration of TTP biased the cleavage toward position -18, with minor products formed at position -17 (Figure 4.5A). Despite different cleavage patterns, the cleavage efficiencies remained almost unchanged by the ligands (Figure 4.5C). Rates of cleavage in the absence of ligands, in the presence of PFA, or using ddAMP-terminated primer with TTP were determined as $11.3 \pm 1.4\ s^{-1}$, $9.4 \pm 0.8\ s^{-1}$, and $13.4 \pm 3.5\ s^{-1}$ for the fast phase, and $0.4 \pm 0.07\ s^{-1}$, $0.4 \pm 0.06\ s^{-1}$, and $0.6 \pm 0.2\ s^{-1}$ for the slower phase. The data clearly disagreed with the toggling mechanism postulating that the engagement of the substrate at the polymerase active site will lead to the exclusive cleavage at a single position in the RNase H site. Rather multiple cleavage bands appeared which also differed in comparing 25 and 26 nt primers. We also noted secondary, slower cleavage events which we attributed to the conversion of $YD_{25}R_{45}$ to $ED_{25}R_{45}$ (Scheme 4.1 and 4.2), the rate of which also remained the same, suggesting that ligands binding to the polymerase site did not affect the rates of conversion from nonproductive to productive states.

Similar conclusions can be drawn from the data obtained with DNA/RNA hybrids containing 26mer primers although different cleavage patterns were observed (Figure 4.5B and D). The DNA 3'-end-directed RNA cleavage failed to strictly follow the translocation status of RT when 26mer primers were used: in the absence of ligands, cleavage products accumulated at 4 distinct positions (from -18 to -15) instead of two that are characteristic of the pre- and post-translocation stages. In the presence of PFA, positions of cuts reduced to 2, namely, -18 and -17. Whereas the next cognate nucleotide (dCTP) was unable to lock RT-DNA/RNA complexes at the post-translocation stage and yielded two major bands at positions -17 and -16. Consequently, RNA cleavage assays were not an ideal method to study translocational dynamics of RT on DNA/RNA hybrids even if sequence biases are avoided as in this study. Similarly, rates of RNA cleavage were not affected by the ligands, which were $10.2 \pm 1.9\ s^{-1}$, $8.8 \pm 0.7\ s^{-1}$, and $10.3 \pm 1.9\ s^{-1}$ for the fast phase, as well as $0.4 \pm 0.13\ s^{-1}$, $0.3 \pm 0.08\ s^{-1}$, and $0.5 \pm 0.13\ s^{-1}$ for the slower phase (Figure 4.5D).

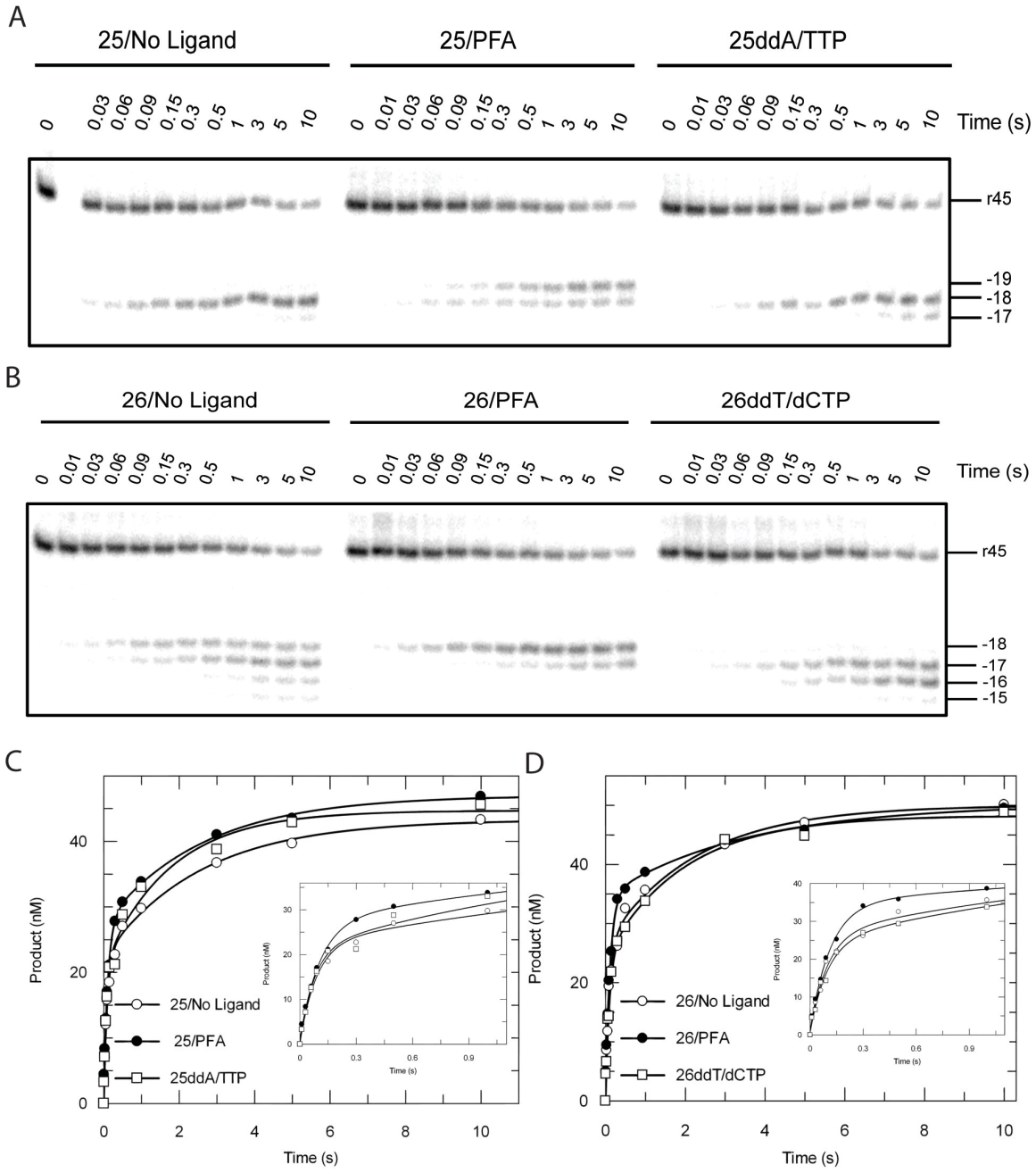

A
25/No Ligand
25/PFA
25ddA/TTP
Time (s)
r45
-19
-18
-17
B
26/No Ligand
26/PFA
26ddT/dCTP
Time (s)
r45
-18
-17
-16
-15
C
Product (nM)
Time (s)
25/No Ligand
25/PFA
25ddA/TTP
D
Product (nM)
Time (s)
26/No Ligand
26/PFA
26ddT/dCTP

**Figure 4.5**

**Figure 4.5** *Effects of Small Ligands on Polymerase-dependent RNase H Activity.* RNA cleavage assays were performed by mixing preformed RT-DNA/RNA complexes (175 nM of HIVRT and 75 nM of d25(ddA)/r45 (A) or d26(ddT)/r45 (B)) with 10 mM $Mg^{2+}$, with or without small ligands. (A) RNase H activities were studied in the absence of any ligands, in the presence of 1 mM PFA, or using ddAMP-terminated primer with next correct nucleotide (0.5 mM TTP). Reactions were stopped by adding of 0.5 M EDTA at various time points up to 10 s. The 45mer RNA template and RNase H cleavages at positions -19 to -17 are indicated. (B) Similar experiments using 26mer primers were performed in the absence of any ligands, in the presence of 1 mM PFA, or using ddTMP-terminated primer with next correct nucleotide (0.5 mM dCTP). The 45mer RNA template and RNase H cleavages at positions -18 to -15 are indicated as well. Cleavage data obtained with d25(ddA)/r45 (C) or d26(ddT)/r45 (D) hybrids were analyzed by conventional data fitting. The insets better display data points up to 1 s. (C) Cleavage data were fit to the double-exponential equation to derive cutting rates measured in the absence of ligands, in the presence of PFA, or using 25ddA with TTP as 11.3 ± 1.4 $s^{-1}$, 9.4 ± 0.8 $s^{-1}$, and 13.4 ± 3.5 $s^{-1}$ for the fast phase, as well as 0.4 ± 0.07 $s^{-1}$, 0.4 ± 0.06 $s^{-1}$, and 0.6 ± 0.2 $s^{-1}$ for the slower phase. (D) Similarly, rates of RNA cleavage using 26mer primers were determined as 10.2 ± 1.9 $s^{-1}$, 8.8 ± 0.7 $s^{-1}$, and 10.3 ± 1.9 $s^{-1}$ for the fast phase, as well as 0.4 ± 0.13 $s^{-1}$, 0.3 ± 0.08 $s^{-1}$, and 0.5 ± 0.13 $s^{-1}$ for the slower phase.

**Effects of Mutations Causing Impaired TTP Induced Conformational Change on RNase H Activity**

Previous studies of HIVRT revealed that the binding of the correct nucleotide could induce large conformational changes in structure from an open to a closed complex (*35, 47*), which was later shown to govern the specificity by HIVRT (*86*). However, whether this crucial conformational change at the polymerase domain affects RNase H activity and/or specificity remains unknown. Here, the activity and specificity of RNA cleavage were examined using three RT mutants (W71A, W71D, and R72A) that were characterized to be able to impair the nucleotide induced isomerization step. The kinetic characterization of the R72A mutant is shown in Chapter 6. Briefly, the kinetic pathway of nucleotide incorporation was quantified by pre-steady-state kinetics together with global data analysis. Then, the extent to which each mutation affected the conformational change step was determined by evaluating the change of the equilibrium constant governing the isomerization step ($K_2 = k_2/k_{-2}$). The results revealed that the R72A mutation had the most adverse effect and impaired the isomerization step by 170-fold. Similarly, W71A and W71D mutations were revealed to disfavor the conformational change step by 12 and 35-fold, respectively (Jiawen Li, unpublished results).

Despite the notable influences on the nucleotide induced conformational change, those mutations had modest or no effect on the activity and specificity of RNA cleavage (Figure 4.6). Similar to the WT protein, all the mutants yielded one major band corresponding to the cutting at position -18 (Figure 4.6A). However, observable cleavage products also accumulated at various positions, including -19, -17, -16 and -15. This was

more evident for W71A and W71D mutants, indicating these two RT mutants were probably less rigidly locked on the DNA/RNA hybrid. In our quantitative analysis, the W71A and W71D mutants showed nearly unchanged cleavage activities compared to the WT protein, with cleavage rates determined as $7.7 \pm 1.3\ s^{-1}$, and $8.1 \pm 0.9\ s^{-1}$ for the faster phase, as well as $0.2 \pm 0.07\ s^{-1}$, and $0.3 \pm 0.04\ s^{-1}$ for the slower phase (Figure 4.6B). The rate of the faster phase for R72A mutant was slightly reduced at $6.3 \pm 1.2\ s^{-1}$. Although the rate of the slower phase remained the same ($0.3 \pm 0.1\ s^{-1}$), its amplitude decreased significantly. As analyzed above (Scheme 4.1 and 4.2), the second slower phase was corresponding to a conversion from an RNase H inactive form ($YD_{25}R_{45}$) to the active form ($ED_{25}R_{45}$), so R72A mutant might impair this process but did not affect the RNase H activity directly. Consequently, the nucleotide induced conformational change might play a role in structural rearrangements involved in conversions between the 'poly-only' and bifunctional species of RT-DNA/RNA complexes but did not directly affect the RNase H activity. These results further supported our 'working simultaneously' model and denied the competing model, in which the two active sites compete with each other to engage substrates and therefore impaired polymerization activity should have led to enhanced RNase H activity.

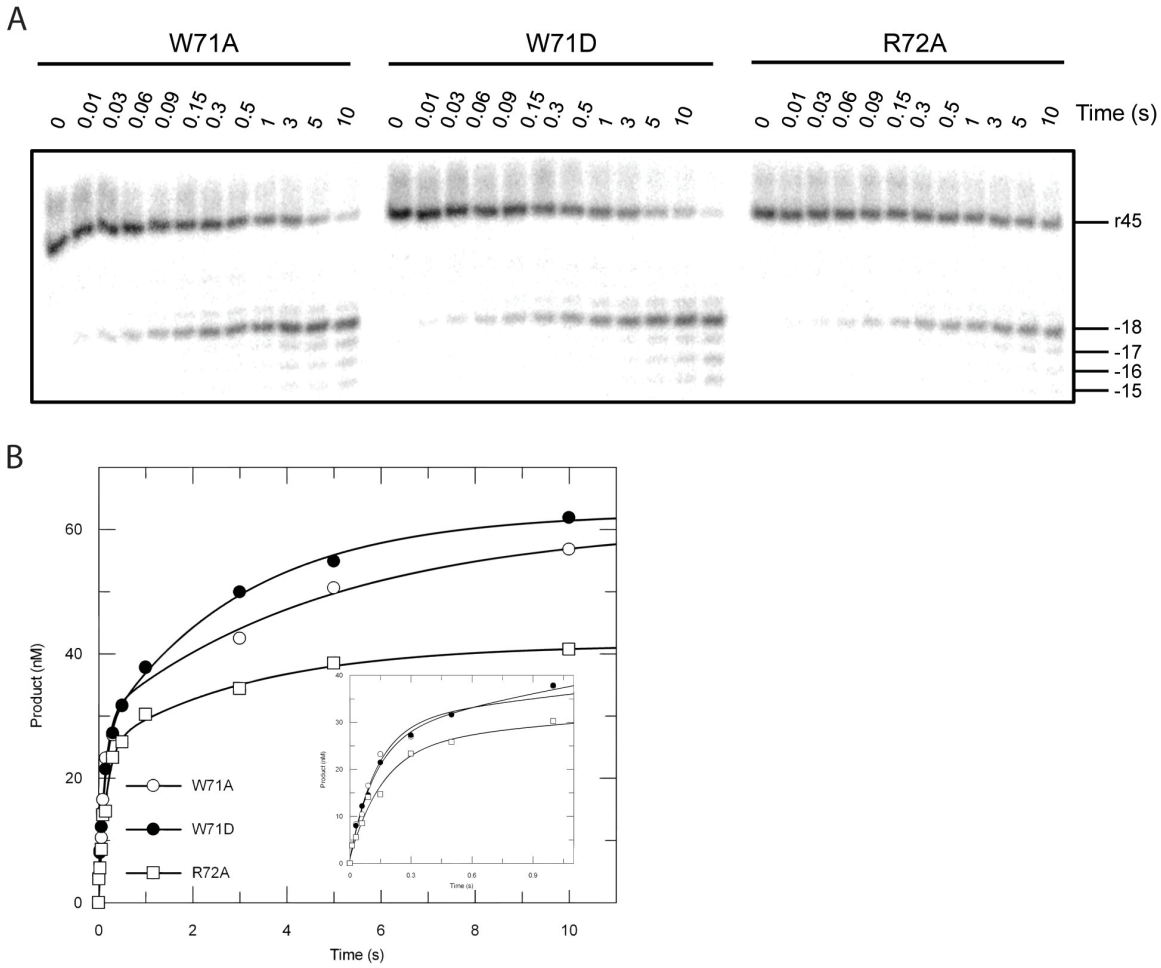



**Figure 4.6** *Effects of Mutations Causing Impaired TTP Induced Conformational Change on RNase H Activity.* (A) RNA cleavage assays were performed by rapidly mixing preformed RT-DNA/RNA complexes (175 nM of HIVRT_W71A, W71D or R72A mutants and 75 nM of d25/r45 hybrid) with 10 mM $Mg^{2+}$. Reactions were stopped by addition of 0.5 M EDTA at various time points in 10 s. The 45mer RNA template and RNase H cleavages at positions -18 to -15 were marked. (B) The Data of RNA cleavage by various RT mutants were analyzed by conventional fitting. The insets better display data points within 1 s. Cleavage data were fit to the double-exponential equation to derive rates of RNA cleavage by W71A, W71D or R72A mutants as 7.7 ± 1.3 $s^{-1}$, 8.1 ± 0.9 $s^{-1}$, and 6.3 ± 1.2 $s^{-1}$ for the fast phase, as well as 0.2 ± 0.07 $s^{-1}$, 0.3 ± 0.04 $s^{-1}$, and 0.3 ± 0.1 $s^{-1}$ for the slower phase.

**Efficient RNase H Cleavage Requires at least 4 Nucleotides downstream of the Cutting Site**

How can the two active sites of HIVRT work simultaneously with different speeds was then examined. Previous studies have suggested that the RNase H activity of HIV RT acted as an endonuclease with partial 3'-5' processivity (*159-160*), which was consistent with our observations that the secondary cuts (38 to 37mer, ~0.05 $s^{-1}$) were 30- to 380-fold slower than the primary cuts (1.5 to 20 $s^{-1}$, Scheme 4.1). Based on these results, a hypothesis was proposed that a minimum length of RNA strand extending from the 3'-end to the cutting site was required for efficient DNA 3'-end-directed RNA cleavages, probably ensuring a proper alignment of the hybrid substrate near the RNase H active site. Under this scenario, the processive RNA cleavage was largely prohibited and a next round of cleavage would occur only when the cutting site shifted towards the 5' end of the RNA strand along with multiple nucleotide incorporation events that could regenerate a downstream RNA strand long enough for the proper alignment of substrates.

In order to test this hypothesis, a series of DNA/RNA hybrids were made by annealing primers with various lengths onto a radioactive-labeled 45mer RNA template with random sequence (Figure 4.7B). RNA cleavage assays were then performed with these hybrids. When the $d25_{rd}/r45_{rd}$ hybrid was examined, two major cleavage products were noticed at position +7 and +8. A new nomenclature of cutting positions were applied here to better illustrate the number of nucleotides extending from the 3' end of the RNA strand to the cleavage site. Theoretically, shortening the primer by one nucleotide will shift the cleavage site accordingly by one nucleotide towards the 3' end of

the RNA strand. The real data, however, did not follow this rule strictly (Figure 4.7A and B). Shortening the primer by 1 to 3 nucleotides all resulted in cleavages at positions +4 and +5, but with different biases. For the $d24_{rd}/r45_{rd}$ hybrid, cleavages products accumulated almost exclusively at position +5, whereas experiments performed with $d23_{rd}/r45_{rd}$ and $d22_{rd}/r45_{rd}$ hybrids showed a strong bias towards +4 cleavage. Further shortening the primer by 1 or 2 nucleotides (21 or 20mer) nearly eliminated the cleavage activity, yielding faint bands at positions +4 and +5, as well as almost negligible bands at position +2. These data supported the idea that at least 4 nucleotides downstream of the cutting site (towards the 3' end of the RNA strand ) were necessary for efficient RNA cleavage to occur. As shown in Figure 4.7C, cleavage data were fit to the single-exponential equation to derive cleavage rates as 5±0.4 $s^{-1}$, 8.8±1.7 $s^{-1}$, 22.8±3.8 $s^{-1}$, 26.7±2.5 $s^{-1}$, 13.4±0.7 $s^{-1}$, and 15.3±2.4 $s^{-1}$ for DNA/RNA hybrids with lengthening primers. Corresponding reaction amplitudes were determined as 16.2±0.4, 18.5±1, 33.5±1.1, 40±1, 33±0.4, and 31.7±1.2, respectively. The quantitative analysis of the cleavage data also revealed that both cleavage rates and amplitudes reduced significantly when the number of nucleotides extending from the 3' end of the RNA strand to the cutting site was less than 4. Consequently, a next round of RNA cleavage would not occur until the incorporation of at least 4 nucleotides, which might account for the fact that the processive template degradation was not observed in previously performed RNA dependent extension assays (*19*).

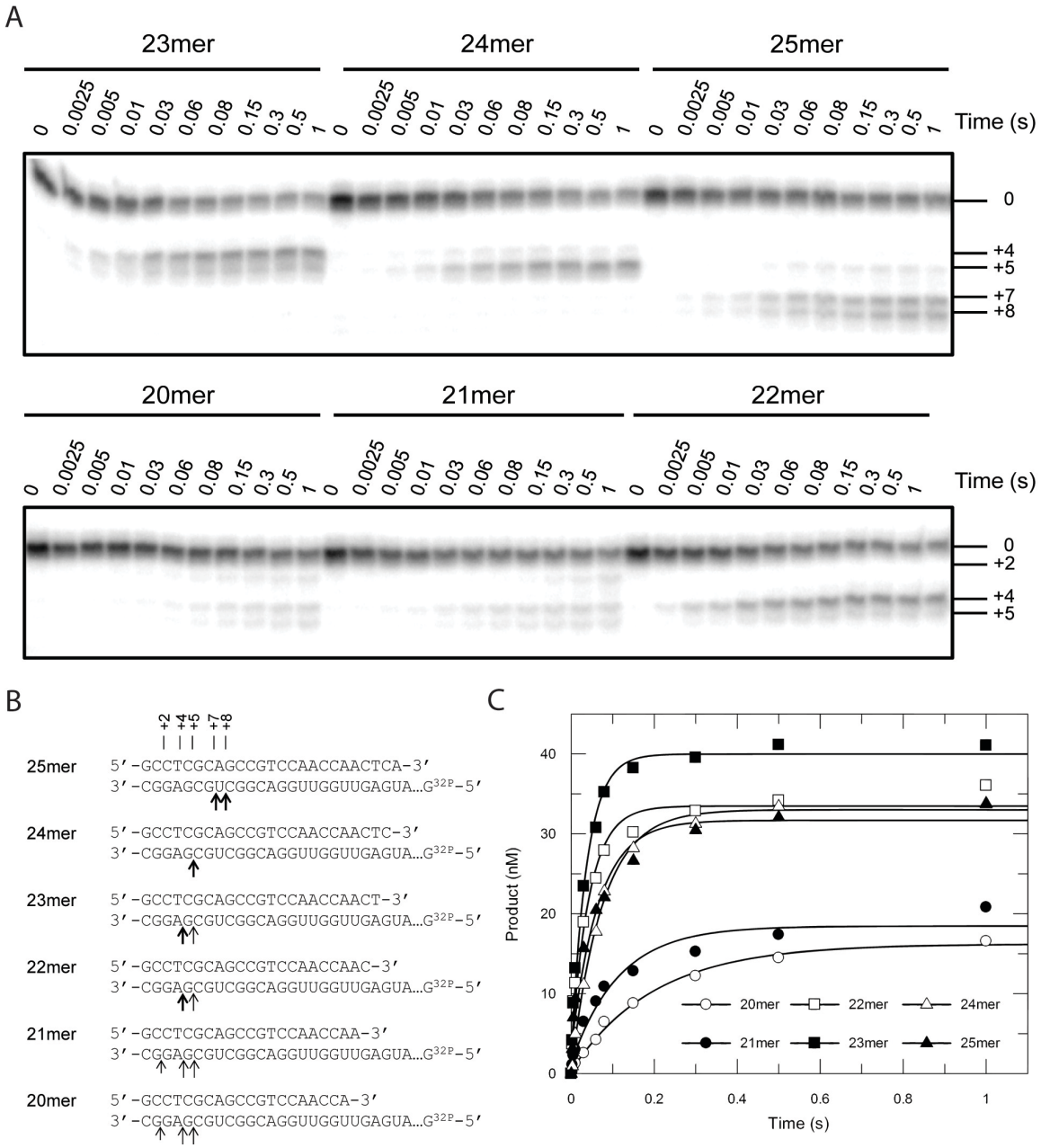

A
23mer
24mer
25mer
0 0.0025 0.005 0.01 0.03 0.06 0.08 0.15 0.3 0.5 1
Time (s)
0
+4
+5
+7
+8
20mer
21mer
22mer
0 0.0025 0.005 0.01 0.03 0.06 0.08 0.15 0.3 0.5 1
Time (s)
0
+2
+4
+5
B
+2 +4 +5 +7 +8
25mer 5′-GCCTCGCAGCCGTCCAACCAACTCA-3′
3′-CGGAGCGUCGGCAGGUUGGUUGAGUA…G32P-5′
24mer 5′-GCCTCGCAGCCGTCCAACCAACTC-3′
3′-CGGAGCGUCGGCAGGUUGGUUGAGUA…G32P-5′
23mer 5′-GCCTCGCAGCCGTCCAACCAACT-3′
3′-CGGAGCGUCGGCAGGUUGGUUGAGUA…G32P-5′
22mer 5′-GCCTCGCAGCCGTCCAACCAAC-3′
3′-CGGAGCGUCGGCAGGUUGGUUGAGUA…G32P-5′
21mer 5′-GCCTCGCAGCCGTCCAACCAA-3′
3′-CGGAGCGUCGGCAGGUUGGUUGAGUA…G32P-5′
20mer 5′-GCCTCGCAGCCGTCCAACCA-3′
3′-CGGAGCGUCGGCAGGUUGGUUGAGUA…G32P-5′
C
Product (nM)
0
10
20
30
40
0
0.2
0.4
0.6
0.8
1
Time (s)
20mer
22mer
24mer
21mer
23mer
25mer

**Figure 4.7**

**Figure 4.7** *Efficient RNase H Cleavage Required at least 4 Ribonucleotides downstream of the Cutting Site.* (A) RNA cleavage assays were performed by rapidly mixing preformed RT-DNA/RNA complexes (175 nM of HIVRT and 75 nM of DNA/RNA hybrids (the 45mer RNA template was annealed with primers of lengths ranging from 20-25 nucleotides) with 10 mM $Mg^{2+}$. Reactions were stopped by adding 0.5 M EDTA at various time points up to 1 s. The position of the uncleaved 45mer RNA template was marked as 0. Different from the nomenclature previously used in this chapter, positions of RNA cleavage products were marked as positive numbers here by counting the number of ribonucleotides from the 3'-end the RNA template to the cutting site. As shown in panel A and B, cleavage sites shifted from +8/+7 to +5/+4 when lengths of primers decreased from 25 to 20 nucleotides. Negligible bands were observed at position +2 in cases of 20 and 21mer. (B) The cleavage sites were illustrated in a cartoon, where thick or light arrows represented cutting sites for major or minor cleavage products, respectively. Short and light arrows indicated the site where negligible cleavage occurred. (D) Cleavage data were fit to the single-exponential equation to derive cleavage rates as 5 ± 0.4 $s^{-1}$, 8.8 ± 1.7 $s^{-1}$, 22.8 ± 3.8 $s^{-1}$, 26.7 ± 2.5 $s^{-1}$, 13.4 ± 0.7 $s^{-1}$, and 15.3 ± 2.4 $s^{-1}$ for DNA/RNA hybrids with lengthening primers. Corresponding reaction amplitudes were determined as 16.2 ± 0.4, 18.5 ± 1, 33.5 ± 1.1, 40 ± 1, 33 ± 0.4, and 31.7 ± 1.2 nM, respectively.

**RNase H Cleavage during Processive Nucleotide Incorporation**

RNase H activities during processive nucleotide incorporation were also explored to further examine the mechanism by which the two catalytic sites of HIVRT were able to work simultaneously but at different rates. As shown in the left panel of Figure 4.8A, the 25mer primer was extended continuously to the full-length 45 nt product, with markers to indicate intermediates with one nucleotide increments. Whereas polymerization is a continuous process at a nearly constant rate, the RNA cleavage activity appears to be discontinuous, supporting our previous hypothesis that at least 4 nucleotides downstream of the cutting site were necessary for efficient RNA cleavage. As we expected, the first round of template degradation started at position +7 with minor bands spreading to position +10 (corresponding to positions -18 to -15 in our previous nomenclature). Incorporation of 9 nucleotides shifted the cutting site accordingly by 9 nucleotides towards the 5' end of the RNA strand, which triggered the second round of RNA cleavage at position +16, with only negligible bands formed between the two rounds of cleavage. Similarly, after the incorporation of 4 and 5 nucleotides, the third and fourth rounds of RNA cleavage started at positions +20 and +25, respectively. On average, the RNA cleavage occurred repeatedly each time after ~6 nucleotides were incorporated.

The incorporation and cleavage data were then fit globally, but in an alternative way due to the complexity of the data, as described in the *materials and methods*. Briefly, the incorporation data were first analyzed according to Scheme 4.3 (Figure 4.8B) to derive the rate constants governing the incorporation of each nucleotide: 94±6, 19±1,

33±3, 30±4, 19±2, 35±8, 36±10, 43±15, 27±6, 29±8, 20±5, 23±7, 36±18, 21±7, 26±12, 21±10, 29±20, 36±34, 20±13, and 10±5 $s^{-1}$, respectively. The average incorporation rate of the 2nd to 9th nucleotides was 27.2±2.7 $s^{-1}$. When the eight-step incorporation was simplified kinetically to a one step process as we did in Scheme 4.4 (+N8), its apparent net incorporation rate was approximately determined by (27.2±2.7)/8 =3.4±0.3 $s^{-1}$. Similarly, apparent incorporation rates for '+N4' and '+N5' were determined as 6.3±1.3 and 5±1.7 $s^{-1}$, respectively.

**Scheme 4.3** *Processive Nucleotides Incorporation*

$$E + D_{25} \rightleftharpoons E(Y)D_{25} \rightleftharpoons XD_{25}$$

$$E(Y)D_{25} + T \rightleftharpoons E(Y)D_{26} \rightleftharpoons E + D_{26}$$

$$\vdots$$

$$E(Y)D_{44} + C \rightleftharpoons E(Y)D_{45} \rightleftharpoons E + D_{45}$$

These rates were then applied to constrain the simulation of RNA degradation data according to Scheme 4.4 (Figure 4.8C). The comprehensive model shown in Scheme 4.4 was derived based on the data and our rule of deriving a minimal model. At the beginning, RT-DNA/RNA complexes were equilibrating between the RNase H active form ($ED_{25}R_{45}$) and inactive forms ($X(Y)D_{25}R_{45}$), with the forward and reverse rates as 0.05 and 0.03 $s^{-1}$. Then, the $ED_{25}R_{45}$ complex underwent RNA cleavage before or after TTP incorporation to yield the 38mer products while the cleaved substrates were still extended. Rates of TTP incorporation (105 $s^{-1}$) and RNA cleavage prior to primer

extension (23 $s^{-1}$) were determined previously (Figure 4.3) and fixed in the simulation to reduce the number of variable parameters. After TTP incorporation, the $ED_{26}R_{45}$ species faced a kinetic partitioning between the RNA cleavage ($ED_{26}R_{38}$) and multiple nucleotide incorporation ($ED_{34}R_{45}$). In order to simplify the model, the incorporation of 8 nucleotides was expressed as only 'one' step (+N8), which resulted in an apparently 'slow' rate of incorporation as 3.4±0.3 $s^{-1}$. Consequently, the rate of RNA cleavage after TTP incorporation was also slower as 4.4 $s^{-1}$, compared with the 13 $s^{-1}$ we determined previously (Figure 4.3). After the incorporation of 8 nucleotides, the second round of RNA cleavage occurred, yielding 29mer cleavage products. The cleavage reactions could start from $ED_{34}R_{45}$ (0.06 $s^{-1}$) or $ED_{34}R_{38}$ (4.8 $s^{-1}$) complexes, but the latter was the kinetically dominant pathway, as indicated by the rapid decrease of 38mer products after an initial accumulation (Figure 4.8C, green curve). Similarly, the cleavage rate determined here (4.8 $s^{-1}$) was constrained by the kinetic partitioning between RNA cleavage and the next round of multiple nucleotide incorporation events (+N4) at an apparent rate of 6.3 ± 1.3 $s^{-1}$. Then, the third round of RNA cleavage could start from three substrates, namely, $ED_{38}R_{45}$ (4.4 $s^{-1}$), $ED_{38}R_{38}$ (0.06 $s^{-1}$), and $ED_{38}R_{29}$ (0.06$s^{-1}$). Since the amount of 29mer cleavage products remained almost constant after its peak value (Figure 4.8C, blue curve), the observed 25mer products was not due to a secondary cut of the 29mer products. Therefore, cleavage starting form $ED_{38}R_{45}$ became the main pathway at this time. The final round of RNA cleavage occurred after the incorporation of another 5 nucleotide (+N5) at a rate of 5 ± 1.7 $s^{-1}$. Cleavage reactions started from 4 precursors, including $ED_{43}R_{45}$ (4.4 $s^{-1}$), $ED_{43}R_{38}$ (0.4 $s^{-1}$), $ED_{43}R_{29}$ (0.06 $s^{-1}$), and $ED_{43}R_{25}$

(0.06 $s^{-1}$). Cleavages of the 45mer RNA precursor ($ED_{43}R_{45}$) remained the dominant kinetic pathway since the amounts of both 29mer and 25mer cleavage products were almost unchanged after their initial formation (Figure 4.8C, blue and yellow curves). Cleavage of the $ED_{43}R_{38}$ precursor corresponded to the slower phase of the disappearance of the 38mer products after the initial rise (Figure 4.8C, green curve) at a rate of 0.4 $s^{-1}$.

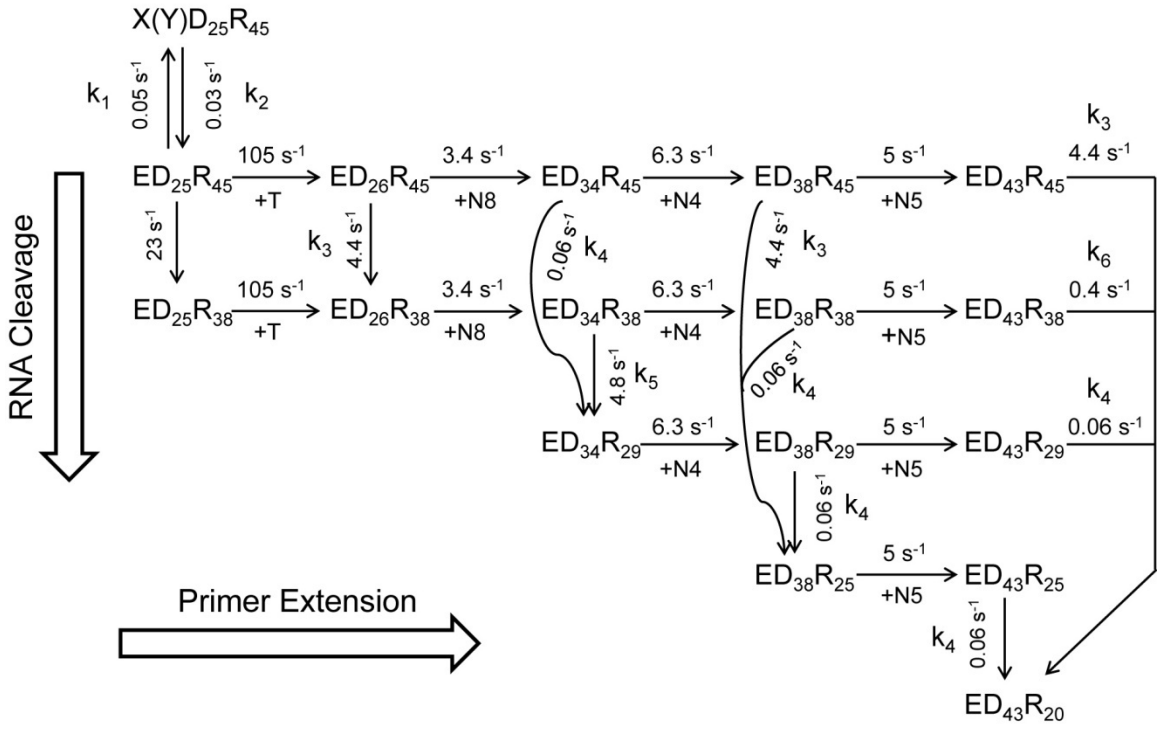
The average rate of nucleotide incorporation was ~28 $s^{-1}$ (except for TTP). However, the 6-fold slower apparent rates were applied in the simulation based on the simplified model. In order to maintain the same kinetic partitioning between RNA cleavage and nucleotide incorporation activities in a complete model, the average rate of RNA cleavage after TTP incorporation could be approximately determined by 4.5×6 = 27 $s^{-1}$. Consequently, under processive nucleotide incorporation conditions, the two catalytic centers of HIVRT could work simultaneously with the same pace. And the discontinuous RNase H activities observed in this study were due to the fact that at least 4 nucleotides downstream of the cutting site were required for efficient RNA cleavage to occur.

**Table 4.3** *Kinetic Parameters Governing RNase H Cleavages during Processive Nucleotide Incorporation*[a]

| $k_1$ ($s^{-1}$) | $k_2$ ($s^{-1}$) | $k_3$ ($s^{-1}$) | $k_4$ ($s^{-1}$) | $k_5$ ($s^{-1}$) | $k_6$ ($s^{-1}$) |
|---|---|---|---|---|---|
| 0.05 | 0.03 | 4.4 | 0.06 | 4.8 | 0.4 |
| (0.043~0.055) | (0.026~0.033) | (3.5~4.9) | (0.05~0.08) | (3.8~6.1) | (0.3~0.6) |

[a] The lower and upper limits of each rate constant derived from confidence contour analysis of the simulation were listed in the parenthesis. The nomenclature of rate constants was illustrated in Scheme 4.4. All parameters were well constrained in the fitting.

**Scheme 4.4** *RNase H Cleavage during Processive Nucleotide Incorporation*



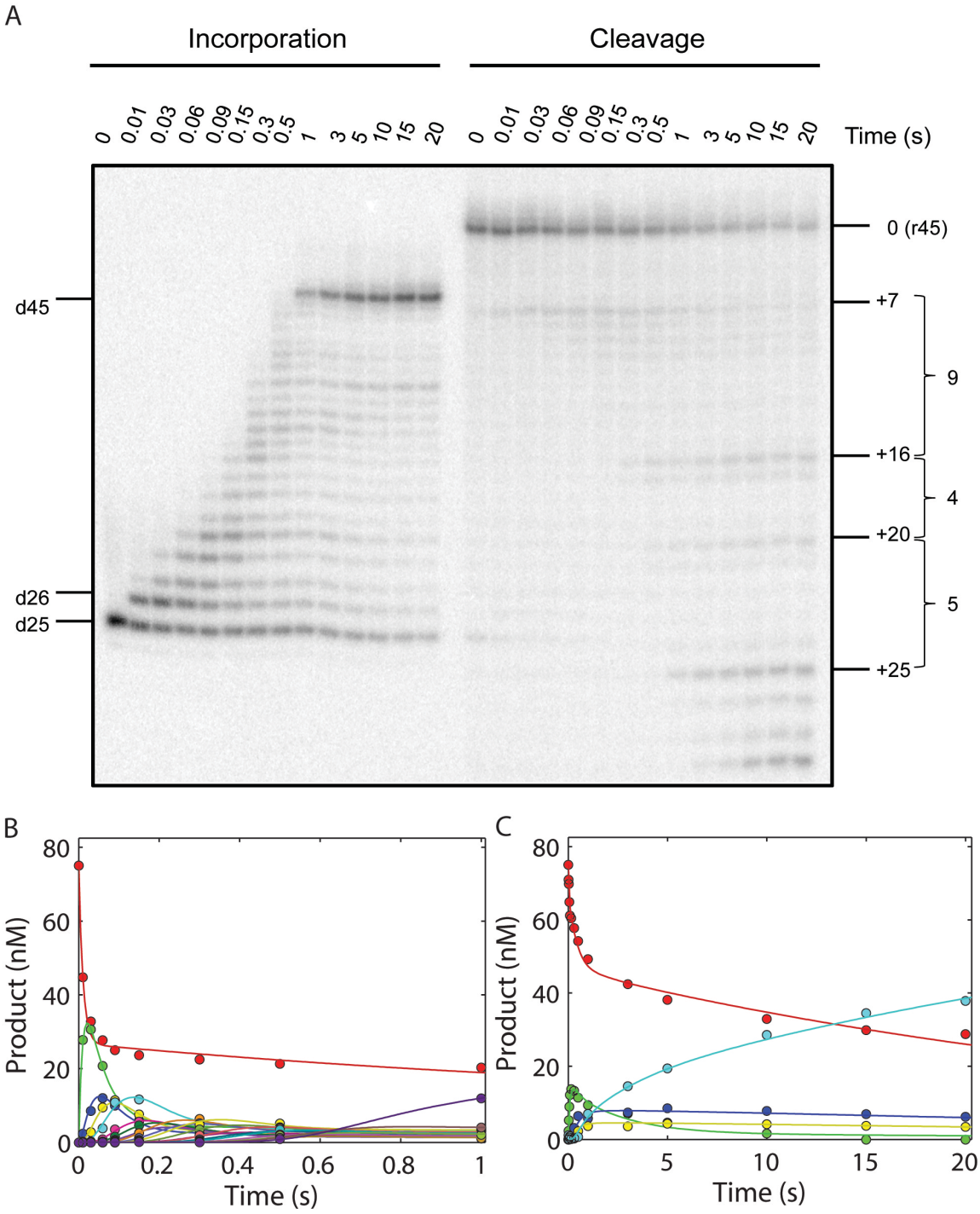

A
Incorporation
Cleavage
0 0.01 0.03 0.06 0.09 0.15 0.3 0.5 1 3 5 10 15 20
0 0.01 0.03 0.06 0.09 0.15 0.3 0.5 1 3 5 10 15 20
Time (s)
d45
d26
d25
0 (r45)
+7
+16
+20
+25
9
4
5
B
C
Product (nM)
Time (s)
Product (nM)
Time (s)

**Figure 4.8**

**Figure 4.8** *RNase H Cleavage during Processive Nucleotide Incorporation.* (A) RNase H activities during processive nucleotide incorporation were explored by performing RNA cleavage assays in the presence of all four nucleotides (100 μM). The procedures were the same as described previously except that data were collected for 20 s to allow full extension and multiple rounds of cleavage to occur. In order to avoid overlaps of bands for extension and cleavage products, the two activities were assayed in parallel as described above. Positions for the primer (d25), primer extended by one nucleotide (d26), and fully extended product (d45) were marked. The same nomenclature as used in Figure 4.8 was applied here to label positions of full length RNA template (0) and cleaved products. Positions for the longest cleavage products formed during each round of degradation were marked as +7, +16, +20, and +25. The numbers of nucleotides incorporated before each round of RNA degradation were also labeled as 9, 4, and 5, respectively. (B) Time courses of processive nucleotide incorporation were simulated according to Scheme 4.3, with smooth lines indicating the best fit. Only data within 1 s were showed to better display the main content of the data and avoid the long non-zero tails of fitting curves. The fitted rates for sequential polymerization steps were 94±6, 19±1, 33±3, 30±4, 19±2, 35±8, 36±10, 43±15, 27±6, 29±8, 20±5, 23±7, 36±18, 21±7, 26±12, 21±10, 29±20, 36±34, 20±13, and 10±5 $s^{-1}$, respectively. The apparent incorporation rates for '+N8', '+N4', and '+N5' were then determined as 3.4±0.3, 6.3±1.3, and 5±1.7 $s^{-1}$, respectively. (C) The apparent incorporation rates obtained from panel B were used in the global analysis of time courses of RNA cleavage according to Scheme 4.4. Apparent incorporation rates and fitted rates for each round of RNA degradation were summarized in Scheme 4.4 and Table 4.3. Color codes: red, 45mer RNA template; green, 1$^{st}$ cleavage (38mer); blue, 2$^{nd}$ cleavage (29mer); yellow, 3$^{rd}$ cleavage (25mer); and cyan, fourth cleavage (20mer).

## 4.4 DISCUSSION

HIV reverse transcription is a complex process that converts the single-stranded RNA genome of the virus into the double-stranded DNA (*161*). This process requires a close coordination of the three functions of HIVRT: DNA- and RNA-dependent polymerase and RNase H, such that the RNase H activity needs to concomitantly degrade the RNA genome during minus-strand DNA synthesis. Therefore, an exploration of the potential coordination between the two active sites of HIVRT is of great importance for a better understanding of reverse transcription. Despite the abundance of structural and biochemical data, whether these two catalytic centers can work simultaneously still eludes researchers in this field. In this study, the RNase H activity during single-nucleotide incorporation was first examined and the results revealed the existence of RT-DNA/RNA complexes that were able to perform primer extension and template degradation activities simultaneously. However, it has been widely accepted that polymerase and RNase H activities coordinate spatially but not temporally since the rate of polymerization is approximately 7-fold faster than RNA cleavage (*20*). The RNase H activity during processive nucleotide incorporation was then examined and the results revealed that the processive polymerization was limited by the slow PPi release (Chapter 5) whose rate was comparable to RNase H cleavage. Under this scenario, the two catalytic centers can coordinate smoothly and harmoniously.

The potential coupling between the two active sites of HIVRT was first explored under single-turnover conditions. Three distinct modes of RNase H cleavage have been described, namely, the DNA 3'-end-directed, RNA 5'-end-directed, and the internal

cleavages, all of which played important roles in the reverse transcription (*18*). And the substrates used in this research were designed to study the first mode of RNA cleavage since the remaining two modes do not involve the polymerization activity. Nucleotide sequence preferences in the vicinity of the cleavage site are another major concern as previous studies indicated that the DNA 3'-end-directed RNA cleavages occurred within a window of ~6 nucleotides centered on the predicted distance (*20-21, 92, 98, 101, 162-163*). In order to remove the sequence biases, the nucleotides near the predicted cleavage sites were replaced by a string of Gs since A tracts are straight and stiff and appear to contribute to the resistance of the PPT to cleavage (*15*).

The incorporation and cleavage data were analyzed by both conventional and global data analysis, and the rate constants obtained from both methods were consistent, supporting the validity of the model chosen to fit the data globally. The only exception was that the rate constant governing the conversion of $XD_{25}R_{45}$ to $YD_{25}R_{45}$ (~0.012 $s^{-1}$, $k_2$ in Scheme 4.1) was much slower than the rate of slower phase (2±0.8 $s^{-1}$) derived from the conventional fitting of the incorporation data (Figure 4.1C) and had a half-time of ~60 s. Consequently, the slower phase of TTP incorporation measured within 10 s might mainly reflect the dissociation and rebinding of the DNA/RNA hybrid substrate (~0.3 $s^{-1}$); this was supported by the fact that the second slower phase almost disappeared when heparin trap was added to prevent rebinding (data not shown). This rationale also explained why individual parameters $k_1$ and $k_2$ (Scheme 4.1) were not well constrained in the confidence contour analysis although their ratio was constant and accurately determined by the simulation (Table 4.2), and a measurement of longer timescale would

help to better define those two rates. During ddTTP incorporation, the corresponding rate of conversion ($XD_{25}R_{45}$ to $YD_{25}R_{45}$) was ~0.04 $s^{-1}$ ($k_2$ in Scheme 4.2) with a shorter half-time (~15 s) and therefore the rate constant was better defined by the data and was well-constrained in the simulation. On the other hand, the slower phase of RNA cleavages (Figure 4.1D) still existed even in the presence of heparin albeit with a slightly reduced amplitude (data not shown), suggesting that the second phase probably involved both the substrate dissociation (and rebinding), as well as the conversion to the RNase H competent species ($YD_{25}R_{45}$ to $ED_{25}R_{45}$). Since the former (~0.3 $s^{-1}$) was defined by the burst experiment (data not shown) and locked in the simulation, parameters governing this conversion were well-constrained in the confidence contour analysis. Actually, most of the parameters were well-defined by our models, as illustrated by the narrow ranges of rate constants listed in Table 4.2. Few exceptions were observed for slow reactions, including the mismatched extension and secondary RNase H cleavages.

Three distinct RT-DNA/RNA complexes were identified in the simulation, including the *bifunctional*, *polymerase active only*, and *inactive* forms. Although active site titration was always performed prior to the assays, 30% or 37% of the RT-DNA/RNA complexes were not active for TTP or ddTTP incorporation in this study. Results of native gel analysis revealed that the DNA primers were in slight excess over the RNA templates in our hybrid substrate, and the numbers reported here were corrected for the interference from free primers. When a DNA/DNA duplex was used, the percentage of nonproductive complexes reduced to ~20% but the 100% activity has never been achieved no matter a DNA or RNA template was to be used. HIVRT was reported

to adopt dynamic binding orientations on the nucleic acid substrates and could switch between these orientations for different functions (*96*), so the inactive complexes identified here might reflect some nonproductive intermediates which have always been observed in quantitative studies on HIVRT polymerization kinetics (*53, 77, 86, 164*).

Extensive interactions between the nucleic acid substrate and the polymerase domain (*35*) of RT would help to tightly hold the primer 3' end at the polymerase active site. Such interactions, however, appear to be fewer in number in the RNase H site domain (*51*), which might allow the substrate to temporarily swing into then out of the RNase H active site, without affecting contacts at the polymerase site (*93*). Consequently, part of the polymerization competent RT-DNA/RNA complexes did not exhibit RNase H activity ($YD_{25}R_{45}$). Around 40% of RT-DNA/RNA complexes were able to perform RNase H cleavage while ~70% were extending the primer. Thus, the sum of the two activities exceeded 100%, providing evidence for the existence of bifunctional enzyme-nucleic acid complexes capable of simultaneous polymerization and RNase H hydrolysis. If the non-productive forms were excluded from the analysis, we computed that >40% of the RT-DNA/RNA complexes could perform both activities simultaneously.

HIVRT on a nucleic acid substrate is in a dynamic equilibrium between the pre- and post-translocation states (*129*). The molecular details of translocation remains elusive but its equilibrium could be adjusted by small-molecule ligands. For example, the PPi analogue PFA may be able to trap the pre-translocation state of HIVRT (*158*) while the next cognate nucleotide could stabilize the post-translocation complex. Previous kinetic studies have shown that the RNase H activity can be inhibited by these two ligands under

steady-state conditions (*93, 95, 162*), indicating that DNA/RNA hybrids stabilized at the polymerase active site were not able to engage the RNase H site simultaneously. However, these steady state results could be explained by the effect of the ligands in slowing the rates of dissociation and rebinding of the DNA/RNA duplex, and a recent study revealed that the binding of PFA did not affect the efficiency of RNase H cleavages under single-turnover conditions (*93*). This issue was revisited in this study using various substrates without nucleotide sequence preferences and the results supported the conclusion that the DNA/RNA hybrid could engage both active sites simultaneously since the ligands capable of stabilizing the polymerase-competent mode did not inhibit the RNase H activity. Although the cleavage patterns were affected by these ligands, they did not follow the predicted modes, especially when the 26mer primers were tested (Figure 4.5B). In the absence of any ligands, RNase H cleavages occurred at 4 distinct positions, ranging from -18 to -15. Since previous studies have suggested that the RNase H activity of HIVRT acted as an endonuclease with partial 3'-5' processivity (*159-160*), cleavages at position -16 and -15 might reflect secondary cuts of the 38/37mer cleavage products (-18/-17). However, the global data analysis revealed that only the 38mer products (-18) underwent negligible level of secondary cleavages (~0.05 $s^{-1}$). Meanwhile, comparable rates were derived for cleavages at positions -18 to -15 (1.5~5 $s^{-1,}$ Scheme 4.1), so the 36/35mer cleavage products were not due to processive, sequential cleavage events; rather they resulted from two independent events. The results were consistent with previous observations that the DNA 3'-end-directed RNA cleavages occurred within a window of ~6 nucleotides centered on the predicted distance (corresponding to position

-20 to -15) (*20-21, 92, 98, 101, 162-163*). Removal of the nucleotide sequence preference only narrowed down the cleavage window to 4 nucleotides, and this low level of specificity is probably due to the lack of a structural element around the RNase H active site required to bind the DNA/RNA hybrid. Therefore, the RNase H site can temporarily slide back and forth on the substrate, which may be achieved through conformational changes of the DNA/RNA hybrid (*51*). Consequently, the RNase H assay is not a reliable method to measure the equilibrium of translocation stages since cleavages did not occur at the fixed distances reflecting the translocation state but rather within a 'cleavage window' that is 4 nucleotides long. In addition, both PFA and the next correct nucleotide failed to lock the enzyme-nucleic acid complex in a fixed translocation state.

The nucleotide-induced conformational change is a crucial step during nucleotide incorporation and was proposed to govern the specificity and analog discrimination by HIVRT (*86*). However, whether this crucial conformational change affects RNase H activity and/or specificity remains unknown. Therefore, we examined how mutants with impaired rate and equilibrium constants governing the conformational change would influence RNase H cleavages. W71 was chosen to make the mutations since MD simulations of the conformational change step (*58*) suggested that this residue was an important contributor to the rate and free energy of the isomerization step. R72 interacts with the PPi moiety of the incoming nucleotide and the arginine to alanine substitution was reported to significantly impede the PPi release step (*37*). These three mutations have not been observed in NRTIs resistance-associated mutations, probably because they all play an import role in polymerization and therefore have high fitness costs. Our results

revealed that these mutations impeded the nucleotide-dependent conformational change step by 12, 35 and 170-fold, respectively. However, they did not affect the RNase H activity, although minor effects on the cleavage pattern were observed. R72A significantly reduced the amplitude of the slower phase during RNA cleavage, presumably by inhibiting substrate dissociation and/or the conversion to the RNase H active form ($YD_{25}R_{45}$ to $ED_{25}R_{45}$). These results further supported the 'working simultaneously' model and argued against the competing model, in which the two active sites compete with each other to engage substrates such that impaired polymerization activity should have led to enhanced RNase H activity or *vice versa*.

Our data have suggested that the two catalytic centers of HIVRT work simultaneously and independently from each other, except for the obvious fact that continued polymerization provides new substrates for the RNase H active site. Our model also provides a possible explanation for the observation that mixing mutants deficient in either site could fully restore reverse transcription (*95*). However, substitutions at W266 and F61 in HIVRT was able to affect the cleavage specificity and render the RNase H unable to generate/remove PPT primer (*165-167*), implying that the polymerase domain might directly involve in the RNase H specificity.

The rate of polymerization was initially shown to be ~7-fold faster than the RNase H cleavage (*20*) and it was shown that the template was not degraded processively in RNA dependent extension assays (*19*). These observations provided the cornerstone of arguments against the 'working simultaneously' model. Consequently, the coordination between the two active sites was also explored under the processive synthesis condition

by pre-steady-state kinetics with global data analysis. Due to the complexity of the model, reactions with comparable rates were linked in the simulation to reduce the number of parameters to be determined. For instance, rates of RNA cleavage starting from 45mer precursors ($ED_{26}R_{45}$, $ED_{38}R_{45}$, and $ED_{43}R_{45}$, except $ED_{34}R_{45}$) were linked in the simulation. Similarly, minor reactions at each round of RNA cleavage were also linked (except $ED_{43}R_{38}$) and yielded a rate of 0.06 $s^{-1}$. The biphasic decrease of 38mer products after their initial rise defined the rates governing the cleavage of $ED_{34}R_{38}$ and $ED_{43}R_{38}$ as 4.8 and 0.4 $s^{-1}$, respectively. Consequently, the simulation only needed to define 6 parameters despite the complexity of the model, which turned out to be all well-constrained (Table 4.3). Since an enormous number of reactions occurred in the process, a simplified model had to be derived for the simulation. For example, multiple incorporation events between sequential rounds of RNA cleavage were 'reduced' to one kinetic step. This simplification, however, resulted in apparently slow rate of nucleotide incorporation, which could be problematic without a proper interpretation of the data. First, the 'real' average rate for polymerization approximately equals the apparent average rate of nucleotide incorporation (~4.7 $s^{-1}$) multiplied by the average number of nucleotides incorporated (~6), which is ~28 $s^{-1}$. Secondly, all non-dead-end species in the model face a kinetic partitioning between the polymerization and RNase H cleavage activities (e.g., $ED_{26}R_{45}$) and the 'real' polymerization rate is 6-fold faster than the apparent rate applied in the simulation. In order to maintain the same kinetic partitioning in a complete model, the average rate of RNase H cleavage derived from this simplified model (~4.5 $s^{-1}$) also needs to be multiplied by 6 to approximately yield the 'real' rate of RNase H cleavage to

be ~27 $s^{-1}$. Consequently, rates of polymerization and RNase H cleavage were comparable to each other. Our data now suggest a model in which processive polymerization at ~28 $s^{-1}$ is coincident with RNase H activity also occurring at ~28 $s^{-1}$ but only after every ~6 nucleotides. Thus, although the RNase H activity appears to be slower than polymerization, the rate of hydrolysis at each site is identical to the rate of polymerization. Accordingly, the RNase H produces products about 6 nucleotides in length, but does not impede the processive motion of the polymerase.

Previous studies have noted that in single turnover experiments, polymerization with an RNA template is much faster than with a DNA template (*20*). Here we noted that the incorporation of the first nucleotide (~105 $s^{-1}$) was much faster than subsequent events. These data suggested that the processive nucleotide incorporation was constrained by a rate-limiting step occurred after the chemistry step for the first nucleotide but before the binding of the next nucleotide. Further studies supported this rate-limiting step was the slow PPi release (Chapter 5). Starting from the second nucleotide, polymerization was no longer faster than RNase H cleavage, therefore these two activities could coordinate spatially and temporally. Under this circumstance, relative levels of these two activities were determined by their kinetic partitioning at each branch point. Therefore, the RNase H cleavage tended to occur when the extension paused (*94*), not because the two activities were mutually exclusive, but because the kinetic partitioning strongly favored the cleavage activity when polymerization stalled. However, we also observed that template degradation was not processive, although polymerization was as fast as the RNase H cleavage. This is consistent with previous observations that the RNase H

activity of HIVRT acted as an endonuclease with only partial 3'-5' processivity (*159-160*).

A recent crystal structure of HIVRT in complex with a DNA/RNA hybrid provided some structural basis for this periodical RNase H activity (*51*). First, there is no structural element of RT that continuously threads the RNA to the RNase H active site and conformation changes of the DNA/RNA hybrid induced by the sliding of RT along the duplex may be required for facilitating the RNA strand to reach the RNase H site. Secondly, the chelation environment and charge distribution at the RNase H active site changes after the cleavage reaction, which may force the cleaved RNA products to temporarily disengage from the RNase H site (*51*). Moreover, the results suggested that at least 4 nucleotides downstream of the cutting site were required for efficient RNA cleavage and therefore the next round RNase H cleavage would not occur until at least 4 nucleotides were incorporated. The oligonucleotide with a minimum length of 4 nucleotides extending from the 3' end of the RNA strand to the cleavage site may help to thread the RNA strand to the RNase H active site, probably through interactions between the oligonucleotide segment and HIVRT near the RNase H site. These interactions, however, may not establish until the segment is long enough. The results also suggested a possible mechanism for HIVRT to prevent uncontrolled template degradation that may cause premature dissociation of the primer strand and termination of synthesis in advance.

In summary, our results establish a mechanism by which polymerization and RNase H activities are coordinated by working independently at comparable rates but with different substrate specificity requirements. Binding of the duplex at the polymerase

active site affords processive synthesis due to the fast translocation and slow dissociation of the duplex from the enzyme such that the net rate is limited by slow pyrophosphate release. In contrast, the RNase H activity requires ~4 – 6 base pairs extending beyond the active site for optimal binding leading to fast rates of cleavage comparable to the rate of polymerization. Although the RNase H reaction is not very efficient at any one site, multiple opportunities for cleavage afford efficient net cleavage during processive synthesis. The inefficient RNase H reaction may, in fact, be an important feature of the enzyme activity so that the enzyme does not binding tightly to the DNA/RNA duplex at any one site since that would impede the movement of the duplex. Consequently, the polymerase and RNase H activities are seamlessly coordinated without the need for any direct communication between the two sites, other than the continuous movement of the duplex due to processive synthesis.

# Chapter 5: Rate-limiting Pyrophosphate Release by HIV Reverse Transcriptase and Its Effect on Enzyme Specificity

## 5.1 INTRODUCTION

HIV-1 Reverse Transcriptase (HIVRT) is a DNA/RNA-dependent DNA polymerase that plays a central role during viral replication. It contains the polymerase and RNase H active sites that are ~60 Å away from each other. In Chapter 4, a rate-limiting PPi release step was identified, which plays an important role in lowering the rate of processive synthesis and synchronizing the two catalytic centers during reverse transcription. However, questions remain as to how it would affect the kinetics of sequential single nucleotide incorporation events during processive synthesis. The slow PPi release has been observed in previous studies of human mitochondrial DNA polymerase when incorporating modified or mismatched nucleotides (*89-91*), where it served to reduce the specificity constant. In this chapter, we examine how the slow PPi dissociation affects the nucleotide selectivity by HIV-1 RT.

Previous studies have demonstrated that a nucleotide-induced conformation change governs the HIV-1 RT specificity (*86*). Measurement of the rates of polymerization and pyrophosphate release in single turnover experiments with DNA templates showed that pyrophosphate (PPi) dissociation was fast following nucleotide incorporation so that it did not contribute to enzyme specificity ($k_{cat}/K_m$). In the current study, kinetic parameters governing nucleotide incorporation and PPi release were determined using an RNA template.

HIV-1 RT has been an important target for antiretroviral therapy due to its critical roles in reverse transcription (*60-61*). Nowadays, drug resistance associated mutations in the RT gene have been identified for all NRTIs approved by the US FDA (*67*). The low fidelity of HIV-1 RT has been proposed to make HIV-1 replication particularly error-prone (*168-169*), allowing the virus to develop resistance to antiretroviral drugs rapidly (*170*). Indeed, the fidelity of HIV-1 RT has been reported to be 10-fold lower than the RTs from other origins (*171-174*). However, the true fidelity of HIV-1 RT and its relative contribution to the high sequence variation of HIV are still being debated. Fidelity measured *in vitro* using purified HIV-1 RT vary widely with different conditions (*20, 52, 169, 175-179*), which, however, are typically in the $10^{-4}$ range (per nucleotide addition). But *in vivo* studies have reported mutation frequencies ranging from $1.4\times10^{-5}$ to $\sim4\times10^{-5}$ per nucleotide per replication cycle (*27-29*), which are 10- to 20-fold lower than the error rate of HIV-1 RT measured *in vitro*. Recently, Achuthan *et al.(180)* claimed that they reconciled, at least in part, these discrepancies by showing that HIV-1 RT is not more error prone than other RTs at physiological concentration of $Mg^{2+}$ (0.25 mM). However, the authors only emphasized a particular mis-incorporation (C:A) and their kinetic analysis was flawed, while more comprehensive data from this lab has revealed an opposite trend where lower [$Mg^{2+}$] resulted in reduced fidelity (Gong and Johnson, unpublished results). Previous data from this lab have shown that slow PPi release provided a novel type of proofreading mechanism by allowing for the reversal of chemistry (*89-90*), but this occurred only in rare cases. In this study, the kinetics of

incorrect nucleotide incorporation were further examined to evaluate if slow PPi release could bring the *in vitro* results in closer agreement with the *in vivo* data.

## 5.2 MATERIALS AND METHODS

### DNA and RNA Substrates for Kinetic Studies

DNA primers and templates (Table 5.1) were ordered from Integrated DNA Technologies (IDT) and purified by gel extraction. The 45 nt RNA templates with semi-random sequence (r45) or containing a string of Gs ($r45_{pG}$) were ordered from IDT with RNase free HPLC purification. Dideoxy-terminated DNA primer ($d25_{dd}$) was made through enzymatic synthesis using RT and purified by gel extraction. Primers used in quench flow assays were 5'-$^{32}$P -labeled with T4 polynucleotide kinase (NEB). Annealing of primers and templates was carried out by mixing the two oligonucleotides at a 1:1.2 molar ratio (excess template) and incubating at 95°C for 5 min (DNA primers and RNA templates were incubated at 67°C for 8 min), followed by slow cooling to room temperature.

### Quench Flow Kinetic Assays

Kinetics of nucleotide incorporations were measured by rapidly mixing a pre-formed enzyme-p/t complex (175 nM WT_HIVRT and 75 nM primer/template) with various concentrations of nucleotide using a KinTek RQF-3 (KinTek Corp.). Substrates (nucleic acids and nucleotides) and RT proteins (labeled or unlabeled) may vary as specified in corresponding text and figure legends. Kinetics of pyrophosphorolysis were measured by hand mixing methods in which a pre-formed enzyme-DNA/RNA complex (100 nM MDCC-labeled WT_HIVRT and 150 nM d26/r45) was mixed with various

concentrations of PPi. Processive nucleotide incorporation was measured by rapidly mixing a pre-incubated enzyme-p/t complex (175 nM WT_HIVRT and 75 nM primer/template) with 100 μM of each nucleotide as specified in the text and corresponding figure legends.

**Stopped Flow Kinetic Assays**

Stopped flow assays were performed to measure the kinetics of a single turnover reactions to define the rates of TTP binding and incorporation. Various concentrations of TTP were rapidly mixed with a pre-formed enzyme-DNA/RNA complex (100 nM MDCC-labeled WT_HIVRT and 150 nM d25/r45) and the reactions were monitored by stopped-flow fluorescence methods. Similar experiments were performed at lower temperatures (5, 10, 15, 20, and 25 °C) and the measured rates were extrapolated to estimate the rate of fingers closure at 37°C, which was too fast to measure directly. To determine the nucleotide dissociation rate, 200 nM enzyme-$DNA_{dd}$/RNA (200 nM MDCC-labeled WT_HIVRT and 300 nM $d25_{dd}$/r45 (Table 5.1)) was pre-incubated with 2 μM of TTP, then the preformed ternary complex was mixed with 1 μM of unlabeled enzyme-DNA/RNA duplex to serve as a trap for free nucleotide. The change in fluorescence of the labeled enzyme provided the rate of opening and nucleotide release. A coupled fluorescence assay described elsewhere (*125*) was used to measure the kinetics of pyrophosphate (PPi) release. Briefly, the preformed enzyme-DNA/RNA duplex (200 nM of MDCC-labeled WT_HIVRT and 300 nM d25/r45) was rapidly mixed with 1 mM TTP in the presence of 0.6 μM of pyrophosphatase (PPase), including the Pi "mop" containing 1 mM of 7-methylguanosine (7-MEG) and 0.2 IU/mL purine nucleoside

phosphorylase (PNPase), as well as 0.5 µM of fluorescently labeled *E.coli* phosphate binding protein (PBP). Control experiments were also performed using DNA template or unlabeled RT protein under specified conditions.

**Table 5.1** *DNA/RNA Hybrids for Kinetic Assays*

*d25/r45:*

d25: 5′-GCCTCGCAGCCGTCCAACCAACTCA-3′
r45: 3′-CGGAGCGUCGGCAGGUUGGUUGAGU**A**GCAGCUAGGUUACGGCAGG-5′

*d25$_{pC}$/r45$_{pG}$:*

d25$_{pC}$: 5′-GCCTCCCCCCCCCCCAACCAACTCA-3′
r45$_{pG}$: 3′-CGGAGGGGGGGGGGGUUGGUUGAGU**A**GCAGCUAGGUUACGGCAGG-5′

*d25$_{dd}$/r45:*

d25$_{dd}$: 5′-GCCTCGCAGCCGTCCAACCAACTCA$_{dd}$-3′
r45 : 3′-CGGAGCGUCGGCAGGUUGGUUGAGU**A**GCAGCUAGGUUACGGCAGG-5′

*d26/r45:*

d26: 5′-GCCTCGCAGCCGTCCAACCAACTCAT-3′
r45: 3′-CGGAGCGUCGGCAGGUUGGUUGAGUA**G**CAGCUAGGUUACGGCAGG-5′

*d25/d45:*

d25: 5′-GCCTCGCAGCCGTCCAACCAACTCA-3′
d45: 3′-CGGAGCGTCGGCAGGTTGGTTGAGT**A**GCAGCTAGGTTACGGCAGG-5′

*d25/d45$_{tgc}$:*

d25 : 5′-GCCTCGCAGCCGTCCAACCAACTCA-3′
d45$_{tgc}$: 3′-CGGAGCGTCGGCAGGTTGGTTGAGT**TGC**AGCTAGGTTACGGCAGG-5′

*d25/d45$_{gtc}$:*

d25 : 5′-GCCTCGCAGCCGTCCAACCAACTCA-3′
d45$_{gtc}$: 3′-CGGAGCGTCGGCAGGTTGGTTGAGT**GTC**AGCTAGGTTACGGCAGG-5′

RNA template and DNA primers used in this study are shown. The templating nucleotide used in the studies are shown in bold and underscored.

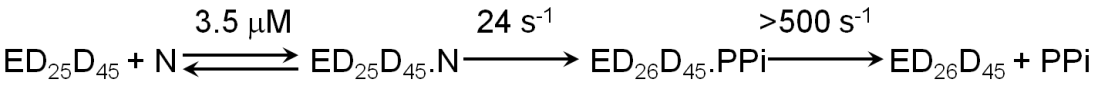
## 5.3 RESULTS

### PPi Release is Fast following DNA polymerization with a DNA Template

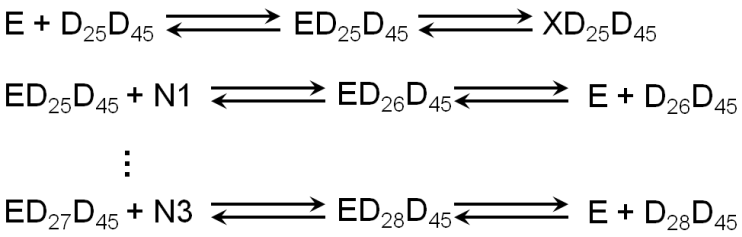
Figure 5.1A shows the time dependence of nucleotide incorporation measured by rapid chemical quench flow methods at several nucleotide concentrations. Figure 5.1B shows the time dependence of the fluorescence change measured by stopped flow during pyrophosphate (PPi) release, coupled to hydrolysis and then binding of phosphate to the fluorescently labeled phosphate binding protein. The data were fit globally, as illustrated by the smooth curves in each figure, to derive the rate constants summarized in Scheme 5.1 and Table 5.3. The apparent binding affinity of TTP ($K_{d,app}$) was 3.5 ± 0.4 μM, and $k_{pol}$ was 24 ± 0.1 $s^{-1}$, yielding a $k_{pol}/K_{d,app}$ value as 6.9 ± 0.8 $\mu M^{-1}s^{-1}$. The time course of PPi release was coincident with the rate of the chemistry step and, accordingly confidence contour analysis set a lower limit of 500 $s^{-1}$ for the rate constant governing PPi release.

Processive nucleotide incorporation assays were performed with two different templates to explore whether there is a significant difference in the rate constants of the first and following nucleotide incorporations (Figure 5.1C-F). The data were fit to a sequential incorporation model as shown in Scheme 5.2, where $XD_{25}D_{45}$ is a nonproductive complex. Rates for the first, second, and third nucleotide incorporations were 23 ± 2, 17 ± 4, and 36 ± 10 $s^{-1}$ for $d25/d45_{tgc}$, as well as 14 ± 1, 26 ± 3, and 30 ± 5 $s^{-1}$ for $d25/d45_{gtc}$. These results contradict the existence of a rate-limiting step after chemistry, which was consistent with the fast PPi release rate (> 500 $s^{-1}$) measured in the single turnover experiment.

**Scheme 5.1** *TTP Incorporation and PPi Release with a DNA Template*




**Scheme 5.2** *Sequential Incorporation Model*



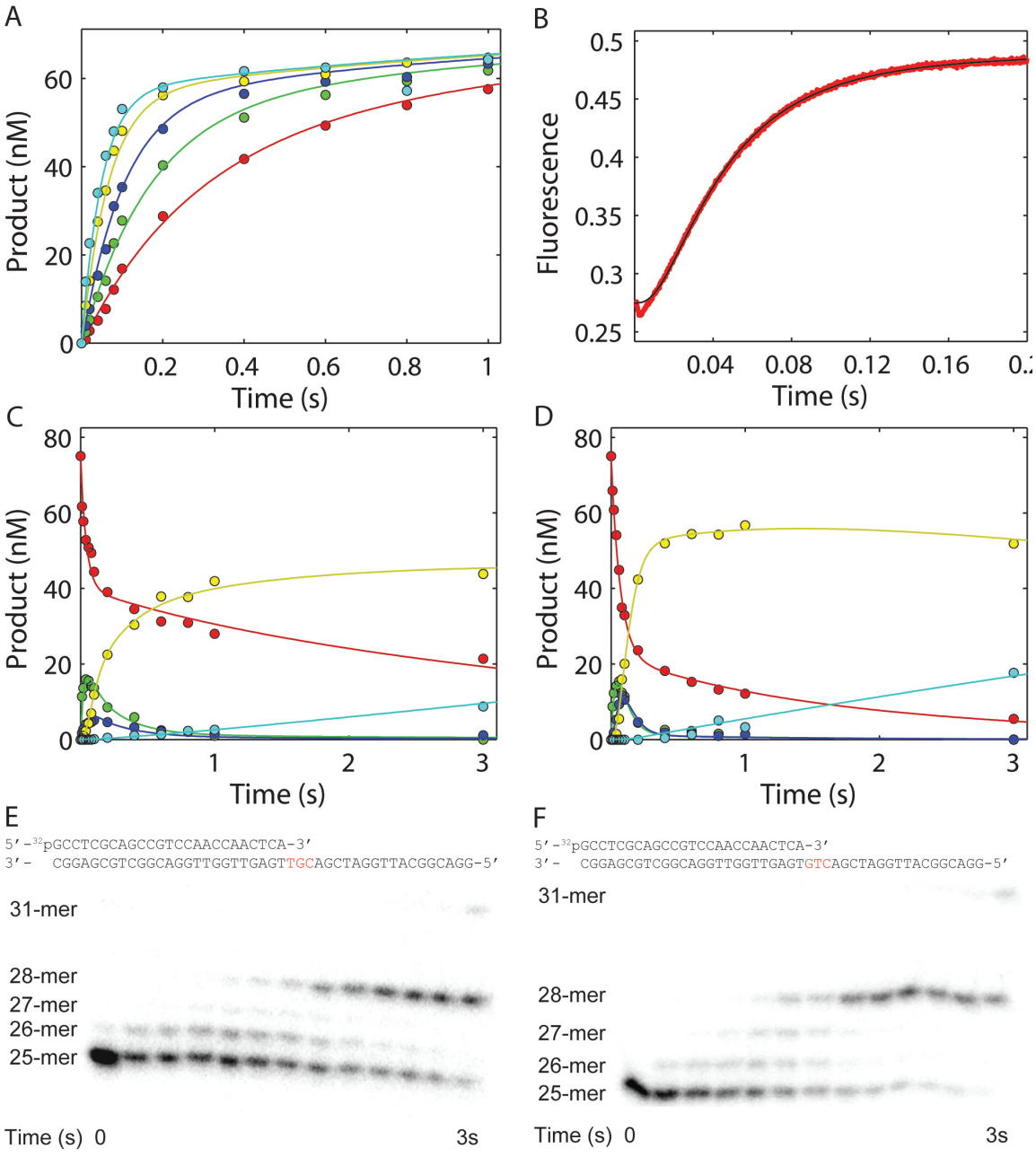

A
Product (nM)
Time (s)
B
Fluorescence
Time (s)
C
Product (nM)
Time (s)
D
Product (nM)
Time (s)
E
5′-32pGCCTCGCAGCCGTCCAACCAACTCA-3′
3′- CGGAGCGTCGGCAGGTTGGTTGAGTTGCAGCTAGGTTACGGCAGG-5′
31-mer
28-mer
27-mer
26-mer
25-mer
Time (s) 0
3s
F
5′-32pGCCTCGCAGCCGTCCAACCAACTCA-3′
3′- CGGAGCGTCGGCAGGTTGGTTGAGTGTCAGCTAGGTTACGGCAGG-5′
31-mer
28-mer
27-mer
26-mer
25-mer
Time (s) 0
3s

**Figure 5.1**

**Figure 5.1** *PPi Release is not a Rate-limiting Step during Nucleotide(s) Incorporation when DNA Templates were used.* (A) Rapid quench flow assays were performed by mixing a pre-formed enzyme-DNA complex (175 nM MDCC-labeled WT_HIVRT and 75 nM d25/d45) with various concentrations of TTP (0.5, 1, 2, 5, and 15 μM). (B) Kinetics of PPi release were measured by mixing a preformed enzyme-DNA duplex (200 nM of MDCC-labeled WT_HIVRT and 300 nM d25/d45) with 1 mM of TTP in the presence of 0.6 μM of pyrophosphatase (PPase), the Pi "mop" containing 1 mM of 7-methylguanosine (7-MEG) and 0.2 IU/mL purine nucleoside phosphorylase (PNPase), and 0.5 μM of fluorescently labeled *E.coli* phosphate binding protein (PBP). Fluorescence was observed by excitation of MDCC at 425 nm and monitoring emission with a 475 nm band-pass filter with a 50 nm bandwidth. (C) Global fitting of processive nucleotide incorporation data is shown in panel E. Rates for the first, second, and third nucleotide incorporation were 23 ± 2, 17 ± 4, and 36 ± 10 $s^{-1}$. (D) Global fitting of sequential processive nucleotide incorporation data shown in panel F. Rates for the first, second, and third nucleotide incorporation were 14 ± 1, 26 ± 3, and 30 ± 5 $s^{-1}$. Color codes: red, 25 nt DNA primer; green, 26 nt product; blue, 27 nt product; yellow, 28 nt product; and cyan, 31 nt mis-incorporation product. (E) Processive nucleotide incorporation was measured by rapidly mixing a pre-incubated enzyme-d25/$d45_{tgc}$ complex (175 nM MDCC-labeled WT_HIVRT and 75 nM d25/ $d45_{tgc}$) with 100 μM of dATP, dCTP, and dGTP. (F) Processive nucleotide incorporation was measured by rapidly mixing a pre-incubated enzyme-d25/$d45_{gtc}$ complex (175 nM MDCC-labeled WT_HIVRT and 75 nM d25/ $d45_{gtc}$) with 100 μM of dATP, dCTP, and dGTP. Positions of 25 nt primer and extension products with different lengths are marked.

**PPi Release Becomes the Rate-limiting Step with a RNA Template**

The kinetics of TTP binding, incorporation, and concomitant PPi release were then examined using an RNA template. Five experiments (Figure 5.2) were conducted and fit simultaneously to a single model as illustrated by Scheme 5.3. Figure 5.2A shows the time dependence of nucleotide incorporation measured by rapid chemical-quench flow methods at several nucleotide concentrations. Figure 5.2B shows the time course of fluorescence changes upon nucleotide binding and incorporation at various nucleotide concentrations. Figure 5.2C shows the time course of fluorescence increase during nucleotide dissociation from the enzyme-$DNA_{dd}$/RNA-TTP ternary complex and then being trapped by unlabeled enzyme-DNA/RNA duplex. Figure 5.2D shows the time dependence of the fluorescence change measured by stopped-flow during pyrophosphate (PPi) release, coupled to hydrolysis and then binding of phosphate to the fluorescent Pi sensor. Figure 5.2E shows the time dependence of pyrophosphorolysis measured by hand quenching at several PPi concentrations. All of the data in Figure 5.2 were fit globally, as illustrated by the smooth curves in each figure, to derive the rate constants summarized in Table 5.2. Among those rate constants, $k_2$ was derived from a global fitting of temperature-dependent TTP incorporation data (Figure 5.3), which was then fixed in the fitting of all five experiments shown in Figure 5.2.

**Scheme 5.3** *Kinetic Pathway of TTP Incorporation with a RNA Template*

$$ED_{25}R_{45} + N \underset{}{\overset{1/K_1}{\rightleftharpoons}} ED_{25}R_{45}.N \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} FD_{25}R_{45}.N \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} FD_{26}R_{45}.PPi \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} ED_{26}R_{45}+PPi$$

The experiments shown in Figure 5.2B were repeated at lower temperatures to derive the maximum rate of nucleotide-induced conformational change ($k_2$) since it was too fast to be measured at 37°C. Traditionally, the extrapolation was achieved by Arrhenius plot (Chapter 2, Figure 2.3C). Herein, a new feature of KinTek Explorer capable of fitting data using temperture-dependent rate constants was employed to obtain the rate constants at 37°C (Figure 5.3). Briefly, a series of experiments covering a range of temperatures (5, 10, 15, 20, and 25 °C) were performed, then the family of curves were fit simultaneously to a single set of rate constants (Scheme 5.3) at the reference temperature (37°C) and a corresponding set of activation energy terms to cover the range of temperatures. A maximum rate of 3200 ± 90 $s^{-1}$ at 37°C was obtained from the fitting.

The value of $k_2$ was then fixed in the fitting of all five experiments shown in Figure 5.2 such that all remaining rate constants were derived from the global data fitting (Table 5.2). The ground state dissociation constant ($1/K_1$) was determined as 520 ± 5 μM. Next, the rate of TTP release from a pre-formed ternary complex was measured by mixing the closed complex with unlabeled trap (Figure 5.2C). 3'-dideoxy-terminated primer was used in this assay to block the chemistry step. Since the ground state binding of TTP is a fast equilibrium process, re-opening of the ternary complex is the rate-limiting step for TTP dissociation and was defined by the assay. This yielded a $k_{-2}$ value of 3±0.01 $s^{-1}$, which was well constrained in the global fitting according to confidence contour analysis (not shown). The rate of chemistry was 250 ± 5 $s^{-1}$, which was 10-fold faster than the rate obtained with a DNA template (24 ± 0.1 $s^{-1}$, Figure 5.1A and Scheme 5.1). The rate of PPi release was examined using a coupled assay previously developed in

our lab (*125*). Briefly, PPi released from the reaction was hydrolyzed by the pyrophosphatase (PPase), yielding inorganic phosphate, which then bound to the fluorescently labeled *E.coli* phosphate binding protein (PBP) to give the fluorescence signal. The kinetics of PPi hydrolysis and Pi binding to MDCC-PBP were calibrated as previously described (Chapter 2). Interestingly, the fitting afforded a PPi release rate of $58 \pm 0.4\ s^{-1}$ that was ~4-fold slower than the rate of chemistry. Therefore, product release became the rate-limiting step (Figure 5.2D). Thanks to the slow PPi release step, the rate constant governing the reversal of chemistry ($k_{-3}$) was able to be accurately determined as $3 \pm 0.3\ s^{-1}$, by measuring the kinetics of pyrophosphorolysis (Figure 5.2E). The experiment also defined a second order rate constant for PPi rebinding of $1.8 \pm 0.4\ \mu M^{-1}s^{-1}$.

**Table 5.2** *Kinetic Parameters Governing TTP Binding and Incorporation by MDCC labeled WT_HIVRT in complex with a DNA/RNA Hybrid*

| $1/K_1$ ($\mu M$) | $k_2$ ($s^{-1}$) | $k_{-2}$ ($s^{-1}$) | $k_3$ ($s^{-1}$) | $k_{-3}$ ($s^{-1}$) | $k_4$ ($s^{-1}$) | $k_{-4}$ ($\mu M^{-1}s^{-1}$) |
|---|---|---|---|---|---|---|
| 520±5 | 3200±90 | 3±0.01 | 250±5 | 3±0.3 | 58±0.4 | 1.8±0.4 |
| $K_{d,net}$[a] ($\mu M$) | $k_{pol}$[b] ($s^{-1}$) | $K_{d,app}$[b] ($\mu M$) | $k_{pol}/K_{d,app}$[b] ($\mu M^{-1}s^{-1}$) | $k_{cat}$ ($s^{-1}$) | $K_m$ ($\mu M$) | $k_{cat}/K_m$[c] ($\mu M^{-1}s^{-1}$) |
| 0.5±0.01 | 201±6 | 37±3 | 5.4±0.5 | 46±2 | 8.1±0.5 | 5.7±0.2 |

[a] $K_{d,net} = 1/(K_1(1+K_2))$, where $K_2 = k_2/k_{-2}$

[b] The values were obtained from the hyperbolic analysis of TTP concentration dependence of incorporation

[c] $k_{cat}/K_m = k_1k_2k_3k_4/(k_2k_3k_4+k_{-1}(k_{-2}k_{-3}+k_{-2}k_4+k_3k_4))$

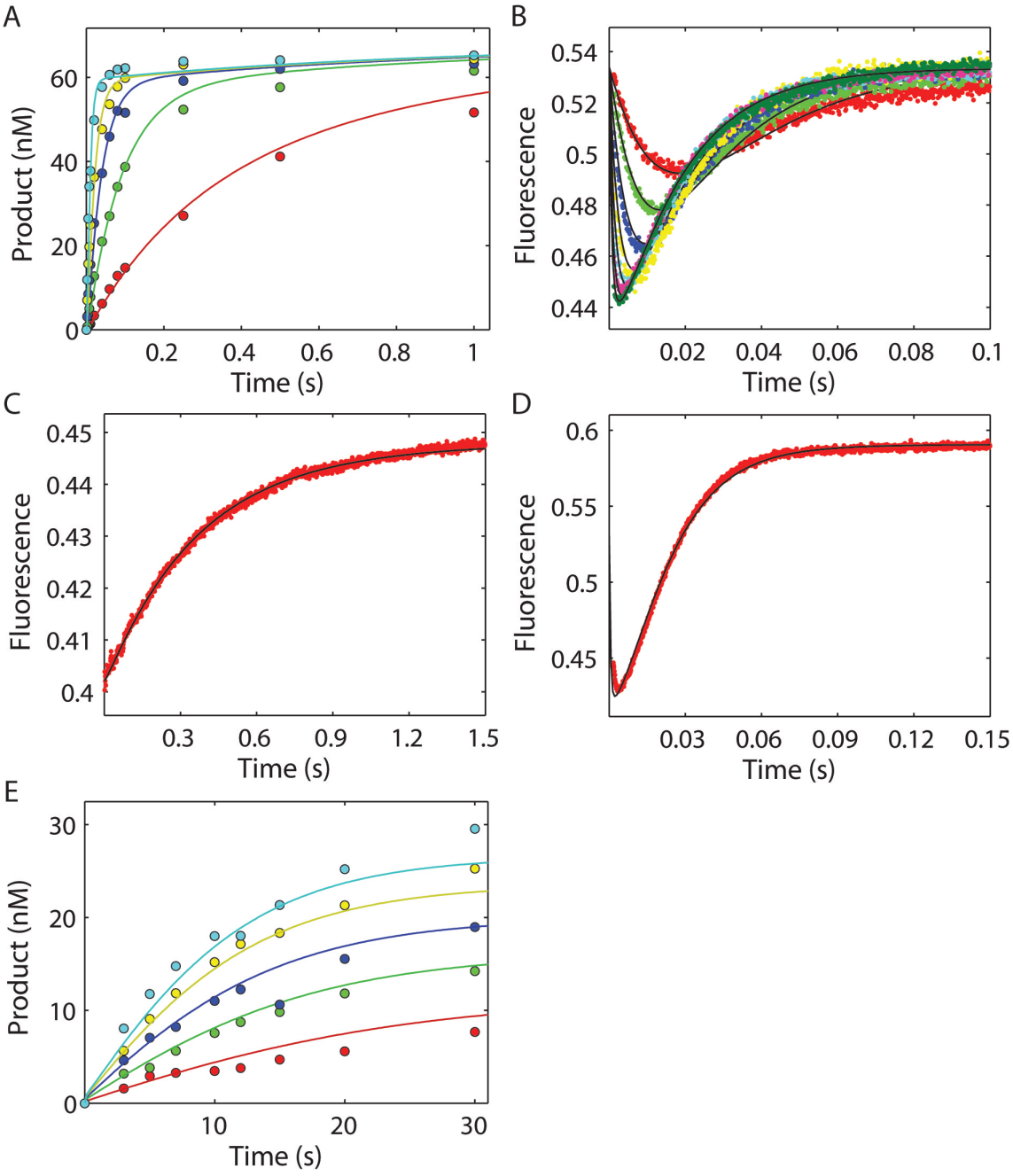

A
Product (nM)
Time (s)
B
Fluorescence
Time (s)
C
Fluorescence
Time (s)
D
Fluorescence
Time (s)
E
Product (nM)
Time (s)

**Figure 5.2**

**Figure 5.2** *Global Fitting of TTP Binding and Incorporation by MDCC-labeled WT_HIVRT using a RNA template*. (A) Rapid quench-flow assays were performed by mixing a pre-formed enzyme-DNA/RNA complex (175 nM MDCC-labeled WT_HIVRT and 75 nM d25/r45) with various concentrations of TTP (0.5, 2, 5, 10, and 30 μM). (B) Stopped flow assays were performed to measure the kinetics of a single turnover of TTP binding and incorporation. Various concentrations of TTP (10, 20, 40, 80, 150, 300, and 600 μM) were rapidly mixed with a pre-formed enzyme-DNA/RNA complex (100 nM MDCC-labeled WT_HIVRT and 150 nM d25/r45) for 100 ms. (C) To determine the nucleotide dissociation rate, 200 nM enzyme-$DNA_{dd}$/RNA (200 nM MDCC-labeled WT_HIVRT, and 300 nM $d25_{dd}$/r45) was pre-incubated with 2 μM of TTP. The preformed ternary complex was then mixed with 1 μM of unlabeled enzyme-DNA/RNA duplex to serve as a trap for free nucleotide. (D) A coupled fluorescence assay (*125*) was used to measure the kinetics of pyrophosphate (PPi) release. Briefly, the preformed enzyme-DNA/RNA duplex (200 nM of MDCC-labeled WT_HIVRT and 300 nM d25/r45) was rapidly mixed with 1 mM of TTP in the presence of 0.6 μM of pyrophosphatase (PPase), the Pi "mop" containing 1 mM of 7-methylguanosine (7-MEG) and 0.2 IU/mL purine nucleoside phosphorylase (PNPase), as well as 0.5 μM of fluorescently labeled *E.coli* phosphate binding protein (PBP). Fluorescence was observed by excitation of MDCC at 425 nm and monitoring emission with a 475 nm band-pass filter with a 50 nm bandwidth. (E) Kinetics of pyrophosphorolysis were measured by hand mixing. A pre-formed enzyme-DNA/RNA complex (100 nM MDCC-labeled WT_HIVRT and 150 nM d26/r45) was mixed with various concentrations of PPi (10, 25, 50, 100, and 200 μM). Global data analysis of all five experiments was performed to derive rate constants governing TTP binding and incorporation, as summarized in Table 5.2 and to generate the smooth curves shown in each panel.

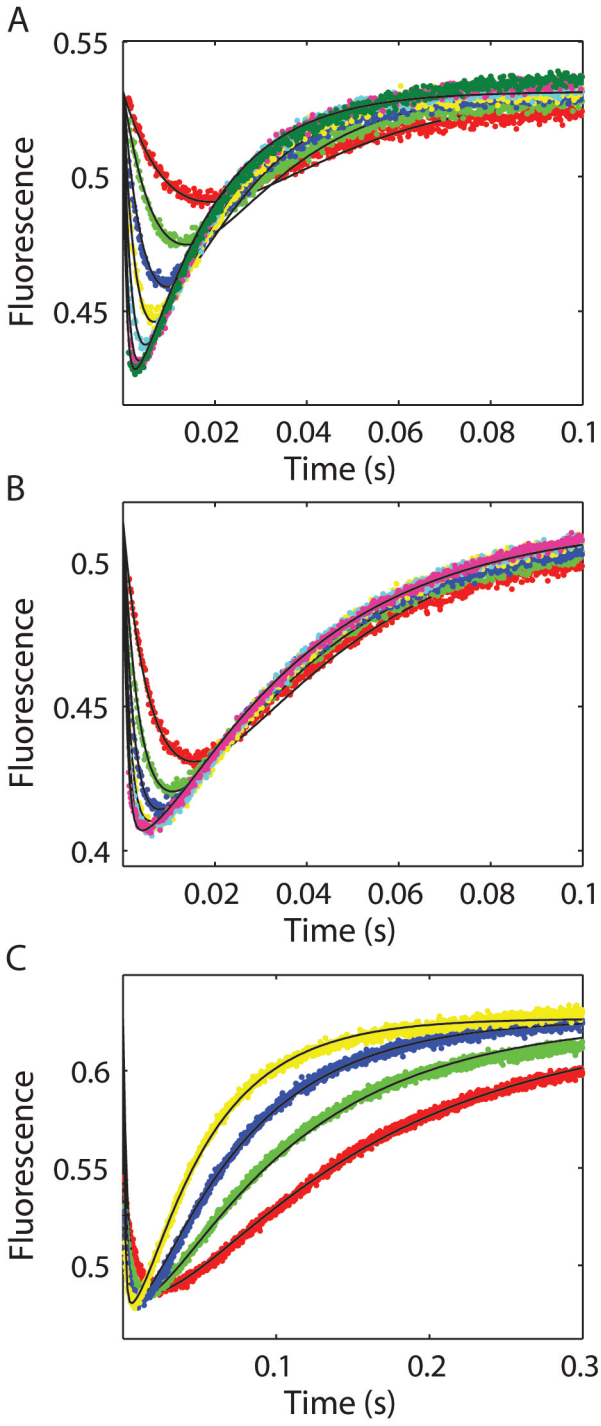



**Figure 5.3** *Global Fitting of Temperature Dependence of TTP Binding and Incorporation by MDCC-labeled WT_HIVRT using a RNA Template.* (A) The time dependence of the fluorescence change was monitored after mixing various concentrations of TTP (10, 20, 40, 80, 150, 300, and 600 μM) with a pre-formed enzyme-DNA/RNA complex (100 nM MDCC-labeled WT_HIVRT and 150 nM d25/r45) at 37°C for 100 ms. (B) Various concentrations of TTP (40, 80, 150, 300, 600, and 1200 μM) were rapidly mixed with a pre-formed enzyme-DNA/RNA complex (100 nM MDCC-labeled WT_HIVRT and 150 nM d25/r45) at 25°C for 100 ms. (C) A preformed enzyme-DNA/RNA complex (100 nM MDCC-labeled WT_HIVRT and 150 nM d25/r45) was rapidly mixed with saturating concentration of TTP (1.5 mM) for 300 ms at 5 (red), 10 (green), 15 (blue), and 20°C (yellow). Fluorescence was observed by excitation of MDCC at 425 nm and monitoring emission with a 475 nm band-pass filter with a 50 nm bandwidth. Global fitting of the experiments performed under six temperatures was performed to derive the rate constant governing TTP induced conformational change ($k_2$) as 3200±90 $s^{-1}$ at 37°C.

Processive nucleotide incorporation was examined to further examine the possible consequences of the rate-limiting PPi release step (Figure 5.4). Incorporation assays were performed with a 25- or 26-nt primer (Figure 5.4 C-D, top). To limit the reaction only three nucleotides were added (no dATP) so the longest extension product was 31 nt, with only negligible formation of a 35 nt mis-incorporation product. The data were fit to a sequential incorporation model similar to the one shown in Scheme 5.2, but with more successive incorporation events. Rates for sequential nucleotide incorporation were 152 ± 9 (T), 18 ± 1 (C), 22 ± 2 (G), 28 ± 5 (T), 23 ± 5 (C), and 29 ± 8 $s^{-1}$ (G) with a 25 nt primer (Figure 5.4A), and 84 ± 5 (C), 28 ± 2 (G), 29 ± 4 (T), 27 ±4 (C), and 32 ± 7 $s^{-1}$ (G) with the 26 nt primer (Figure 5.4B). As shown in Figure 5.4A, there was a large difference in rate constants comparing the first (TTP, 152 ± 9 $s^{-1}$) and following nucleotides incorporation (CGTCG, average rate of 24 ± 10 $s^{-1}$). Similarly, when a 26 nt primer was used, incorporation of the first nucleotide (dCTP, 84 ± 5 $s^{-1}$) was much faster than the following nucleotides (GTCG, average rate of 29 ± 9 $s^{-1}$), suggesting the existence of a rate-limiting step occurring after the formation of the first phosphodiester bond. The conclusion was further supported by the observation that RT incorporated dCTP at a higher rate (84 ± 5 $s^{-1}$) when it was the first incoming nucleotide, whereas at a much slower rate when it was the second nucleotide to be incorporated (18 ± 1 $s^{-1}$) during processive synthesis. Based on these results, the slow PPi release step was identified as the rate-limiting step during processive synthesis.

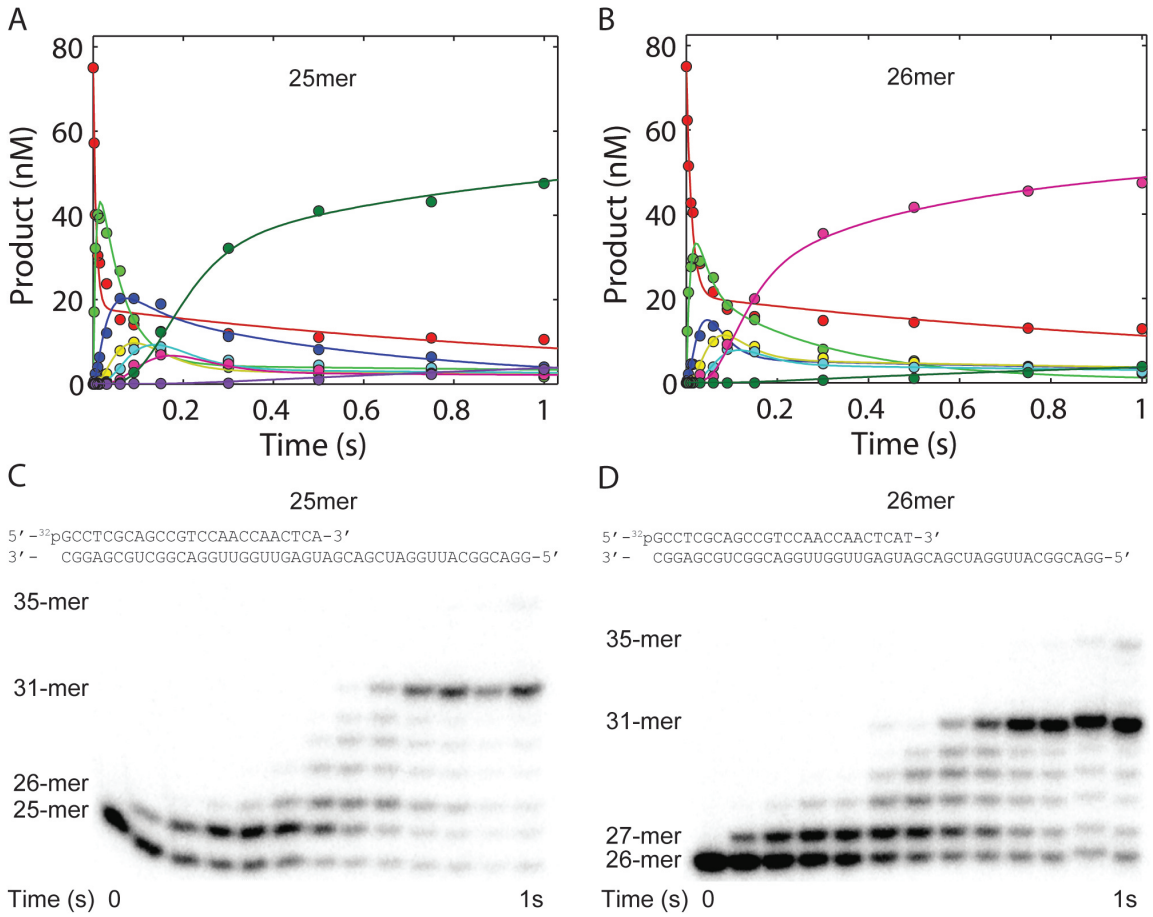



**Figure 5.4** *Slow PPi Release was Supported by Processive Nucleotide Incorporation Assays.* (A) Global data fitting of processive nucleotide incorporation data shown in panel C. Rates for sequential nucleotide incorporation were 152 ± 9, 18 ± 1, 22 ± 2, 28 ± 5, 23 ± 5, and 29 ± 8 $s^{-1}$. Color codes: red, 25 nt DNA primer; light green, 26 nt product; blue, 27 nt; yellow, 28 nt; cyan, 29 nt; magenta, 30 nt; green, 31 nt; and purple, 35 nt mis-incorporation product. (B) Simulation of processive nucleotide incorporation data shown in panel D. Rates for sequential nucleotide incorporation were 84±5, 28±2, 29±4, 27±4, and 32±7 $s^{-1}$. Color codes: red, 26 nt DNA primer; light green, 27 nt product; blue, 28 nt; yellow, 29 nt; cyan, 30 nt; magenta, 31 nt; and green, 35 nt mis-incorporation product. (C-D) Processive nucleotide incorporation was measured by rapidly mixing a pre-incubated enzyme-DNA/RNA complex (175 nM MDCC-labeled WT_HIVRT and 75 nM d25/r45 (C) or d26/r45 (D)) with 100 μM of TTP, dCTP, and dGTP. Sequences of oligonucleotides are shown on the top of each panel. Positions of primers and extension products with different lengths are marked. Timescales are shown on the bottom of each panel.

**Rate-limiting PPi Release Is Not due to the MDCC Labeling**

Previous experiments were all performed with a MDCC labeled WT_HIVRT. Although control experiments were performed to support that both the mutations and the MDCC fluorophore did not affect the incorporation efficiency of HIVRT (*86*), potential effect of MDCC labeling on the PPi release step has not been examined yet. Consequently, kinetics of TTP incorporation and concomitant PPi release by unlabeled WT_HIVRT were also examined with two different DNA/RNA hybrid substrates (Figure 5.5). The nucleotide concentration dependence of the rates of TTP incorporation by the unlabeled WT_HIVRT in complex with a $d25_{pC}/r45_{pG}$ hybrid defined a rate of incorporation ($k_{pol}$) of 150 ± 10 $s^{-1}$, an apparent $K_d$ ($K_{d,app}$) of 34 ± 4 μM, and a specificity constant ($k_{pol}/K_{d,app}$) of 4.4 ± 0.6 $\mu M^{-1}s^{-1}$ (Figure 5.5A). The rate of PPi release was 28 ± 1 $s^{-1}$, as measured by the coupled fluorescence stopped-flow assay (Figure 5.5B). Consequently, PPi release was the rate-limiting step regardless of the presence or absence of the MDCC labeling. Similar experiments and data analysis were performed with a d25/r45 hybrid (Figure 5.5C-D), and the data defined an apparent binding affinity of TTP ($K_{d,app}$) as 24 ± 1 μM, and a rate of chemistry ($k_{pol}$) of 192 ± 8 $s^{-1}$, yielding a $k_{pol}/K_{d,app}$ value as 8 ± 0.5 $\mu M^{-1}s^{-1}$. Again, the rate of PPi release (35 ± 1 $s^{-1}$) was much slower than that of chemistry, demonstrating that the rate-limiting PPi release was not an artifact caused by the MDCC labeling.

**Table 5.3** *Specificity Constants and Rates of PPi release of TTP Incorporation by Various RTs in complex with Different Substrates*

| Enzyme | Substrates | $k_{pol}$ ($s^{-1}$) | $K_{d,app}$ (μM) | $k_{pol}/K_{d,app}$ ($\mu M^{-1}s^{-1}$) | PPi Release ($s^{-1}$) |
|---|---|---|---|---|---|
| MDCC-WT_HIVRT | d25/d45 | 24±0.1 | 3.5±0.4 | 6.9±0.8 | >500 |
| Unlabeled WT_HIVRT | $d25_{pC}/r45_{pG}$ | 150±10 | 34±4 | 4.4±0.6 | 28±1 |
| Unlabeled WT_HIVRT | d25/r45 | 192±8 | 24±1 | 8±0.5 | 35±1 |

Processive nucleotide incorporation assays were also performed using unlabeled WT_HIVRT in complex with two different DNA/RNA hybrids to further support the rate-limiting PPi release step is not due to the MDCC labeling (Figure 5.6). Rates for sequential nucleotide incorporation were 98 ± 6 (T), 22 ± 2 (C), 32 ± 4 (G), 33 ± 5 (T), 21 ± 3 (C), and 27 ± 6 $s^{-1}$ (G) for the $d25_{pC}/r45_{pG}$ hybrid (Figure 5.6A), as well as 92 ± 11 (C), 44 ± 7 (G), 34 ± 6 (T), 23 ± 4 (C), and 29 ± 7 $s^{-1}$ (G) for the $d26_{pC}/r45_{pG}$ hybrid (Figure 5.6B). Again, a significant difference in rate constants of the first and following nucleotides incorporation was observed ($d25_{pC}/r45_{pG}$: 98 ± 6 vs 27 ± 9 $s^{-1}$; and $d26_{pC}/r45_{pG}$: 92 ± 11 vs 32 ± 12 $s^{-1}$), suggesting that the rate-limiting PPi release step occurred after the formation of the first phosphodiester bond during processive syntheis. An additional evidence for this conclusion was that RT incorporated dCTP at a higher rate (92 ± 11 $s^{-1}$) when it was the first incoming nucleotide whereas at a much slower rate when it was the second nucleotide to be incorporated (22 ± 2 $s^{-1}$).

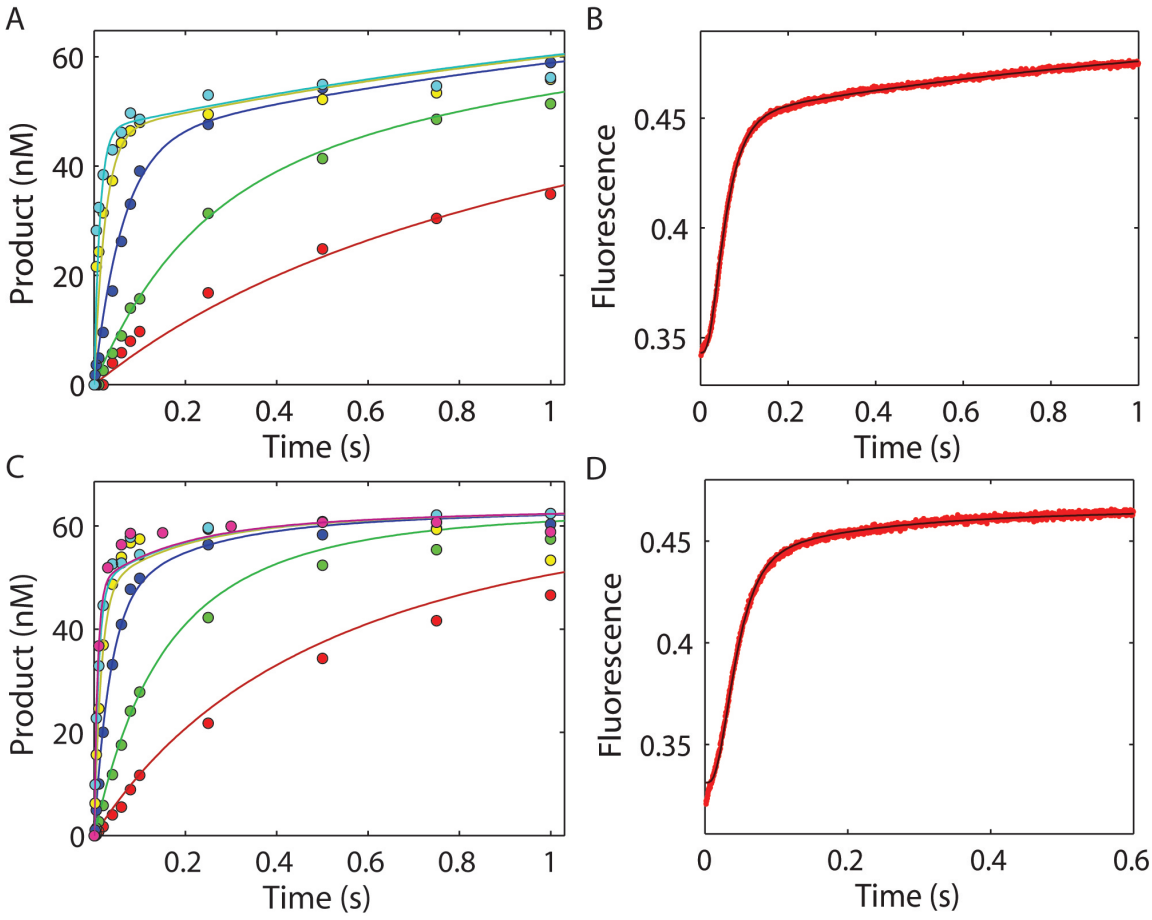



**Figure 5.5** *Global Fitting of TTP Incorporation and Concominant PPi Release By Unlabeled WT_HIVRT.* Rapid quench-flow assays were performed by mixing a pre-formed enzyme-DNA/RNA complex (175 nM unlabeled WT_HIVRT and 75 nM $d25_{pC}/r45_{pG}$ (A) or d25/r45 (C)) with various concentrations of TTP (0.25, 0.75, 3, 10, and 30 µM) (A) or (0.25, 0.75, 3, 10, 30, and 100 µM) (C). Kinetics of PPi release were measured by mixing a preformed enzyme-DNA/RNA duplex (300 nM of unlabeled WT_HIVRT and 200 nM $d25_{pC}/r45_{pG}$ (B) or d25/d45 (D)) with 25 µM of TTP in the presence of 0.6 µM of pyrophosphatase (PPase), the Pi "mop" containing 0.1 mM of 7-methylguanosine (7-MEG) and 0.02 IU/mL purine nucleoside phosphorylase (PNPase), and 0.5 µM of fluorescently labeled *E.coli* phosphate binding protein (PBP). Fluorescence was observed by excitation of MDCC at 425 nm and monitoring emission with a 475 nm band-pass filter with a 50 nm bandwidth. Global analysis was performed independently for $d25_{pC}/r45_{pG}$ (A-B) and d25/r45 (C-D). Rate constants resulting from global analysis of kinetic data are shown in Table 5.3.

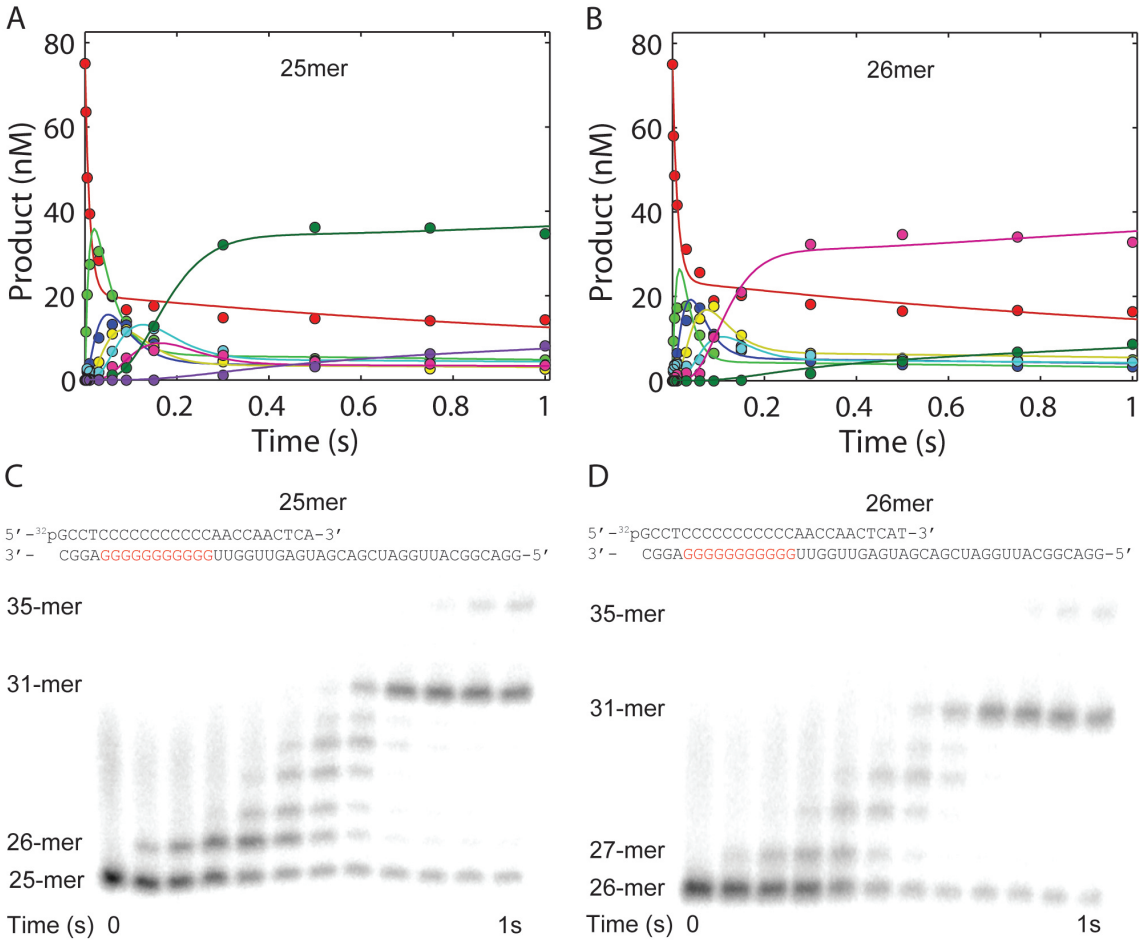



**Figure 5.6** *Processive Nucleotide Incorporation by Unlabeled WT_HIVRT Also Supported Slow PPi Release.* (A) Simulation of processive nucleotide incorporation data shown in panel C. Rates for sequential nucleotide incorporation were 98 ± 6, 22 ± 2, 32 ± 4, 33 ± 5, 21 ± 3, and 27 ± 6 $s^{-1}$. Color codes: red, 25 nt DNA primer; light green, 26 nt product; blue, 27 nt; yellow, 28 nt; cyan, 29 nt; magenta, 30 nt; green, 31 nt; and purple, 35 nt mis-incorporation product. (B) Simulation of processive nucleotide incorporation data shown in panel D. Rates for sequential nucleotide incorporation were 92 ± 11, 44 ± 7, 34 ± 6, 23 ± 4, and 29 ± 7 $s^{-1}$. Color codes: red, 26 nt DNA primer; light green, 27 nt product; blue, 28 nt; yellow, 29 nt; cyan, 30 nt; magenta, 31 nt; and green, 35 nt mis-incorporation product. Processive nucleotide incorporation was measured by rapidly mixing a pre-incubated enzyme-DNA/RNA complex (175 nM unlabeled WT_HIVRT and 75 nM $d25_{pC}/r45_{pG}$ (C) or $d26_{pC}/r45_{pG}$ (D)) with 100 µM of TTP, dCTP, and dGTP. Sequences of substrates are shown on the top of each panel. Positions of primers and extension products with different lengths are marked. Timescales are shown on the bottom of each panel.

**Incorrect Nucleotide Incorporation Also Revealed Rate-limiting PPi Release**

Rapid quench flow assays were performed to measure the kinetics of dGTP:rA mis-incorporation by unlabeled WT_HIVRT in complex with d25/r45 (Figure 5.7A). In order to further explore the slow reactions, the same experiments were repeated for extended timescales, either 5 (Figure 5.7B) or 30 min (Figure 5.7C). Our previous data have shown that the kinetics of mismatch nucleotide binding were complex and most of the fluorescent signal disappeared within the 1 ms dead time of the stopped flow instrument due to the fast opening rate (*58*). Consequently, only quench flow assays were performed and a simplified mechanism involving one-step substrate binding was applied to describe the kinetics of dGTP mis-incorporation (Scheme 5.4). Kinetic parameters resulting from global analysis of kinetic data are summarized in Table 5.4. Data shown in Figure 5.7A defined the rate of mis-incorporation or $k_{pol}$ (approximates $k_2$) as 0.6 ± 0.03 $s^{-1}$ and the apparent nucleotide dissociation constant or $K_{d,app}$ of 11 ± 2 mM. However, a concentration dependence of the amplitude of dGTP incorporation was observed, implying that the chemistry step was reversibly linked to nucleotide binding ($k_{-2}$ = 0.07 ± 0.01 $s^{-1}$) and implying the existence of a slow, rate-limiting step after chemistry (*89-90*). Interestingly, the extension reaction revealed an additional slower phase when dGTP mis-incorporation proceeded for 5 min, where the slow phase was likely to correspond to the rate-limiting PPi release (*89*). Accordingly, the slow phase defined the rate of PPi release to be 0.003 ± 0.001 $s^{-1}$, which was 200-fold slower than the catalytic step and was probably limited by the reopening of RT after primer extension. It is likely that the enzyme remains in a closed conformation after the chemistry step, in which residues

would still be in an alignment to perform catalysis (*89*). This hypothesis is consistent with the fact that the amplitude of the fast phase is dependent on dGTP concentration. But once the enzyme reopens and releases PPi, the reaction would precede to a common end point since PPi release is essentially irreversible due to the relatively low concentration of PPi formed during the course of the reaction (*89*). To test this, kinetics of mis-incorporation were monitored for 30 min (Figure 5.7C). To our surprise, the dependence of reaction end point on nucleotide concentration was still observed, which fit best with a rebinding rate of 0.02 $\mu M^{-1}s^{-1}$. However, any number smaller than this rate could generate an acceptable fitting. Experiments performed under longer timescales could help to better define the rate.

**Scheme 5.4** *Kinetic Pathway of dGTP:dA Mis-incorporation*

$$ED_{25}R_{45} + N \underset{}{\overset{K_{d,app}}{\rightleftharpoons}} FD_{25}R_{45}.N \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} FD_{26}R_{45}.PPi \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} ED_{26}R_{45}+PPi$$

**Table 5.4** *Kinetic Parameters Governing dGTP Binding and Incorporation by Unlabeled WT_HIVRT*

| $K_{d,app}$ (mM) | $k_2$ ($k_{pol}$, $s^{-1}$) | $k_{-2}$ ($s^{-1}$) | $k_3$ ($s^{-1}$) | $k_{-3}$ ($\mu M^{-1}s^{-1}$)[a] |
|---|---|---|---|---|
| 11±2 | 0.6±0.03 | 0.07±0.01 | 0.003±0.001 | 0~0.02 |
| $k_{pol}/K_{d,app}$ ($\mu M^{-1}s^{-1}$) | $k_{cat}$ ($s^{-1}$) | $K_m$ (mM) | $k_{cat}/K_m$[b] ($\mu M^{-1}s^{-1}$) | D[c] |
| $(5.8\pm1.1)\times10^{-5}$ | $(2.5\pm0.6)\times10^{-3}$ | 1.1±0.2 | $(2.4\pm0.7)\times10^{-6}$ | 24±8 |

[a] $k_{-3}$ is not well determined by the data

[b] $k_{cat}/K_m = k_1k_2k_3/(k_2k_3+k_{-1}(k_{-2}+k_3))$

[c]Difference between the two specificity constants obtained by including or overlooking the slow PPi release step was defined by $D = (k_{pol}/K_{d,app}) / (k_{cat}/K_m)$

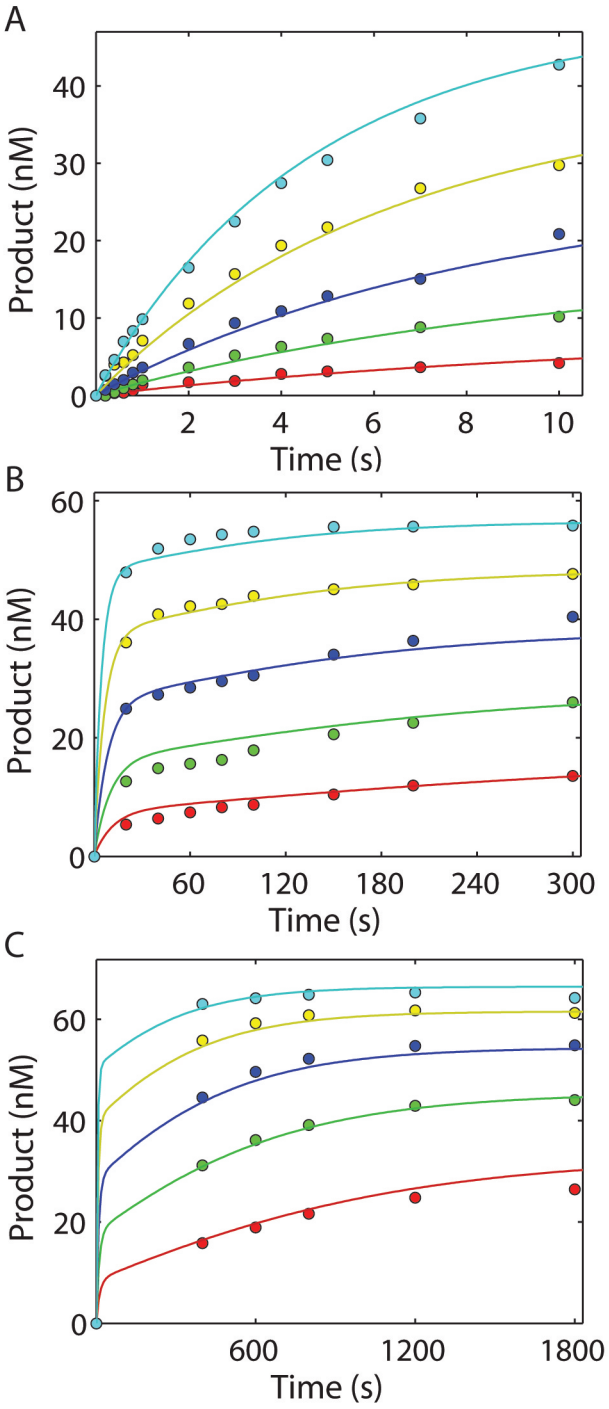



**Figure 5.7** *Global Fitting of dGTP:rA Mis-incorporation by Unlabeled WT_HIVRT.* (A) Rapid quench flow assays were performed by mixing a pre-formed enzyme-DNA/RNA complex (175 nM unlabeled WT_HIVRT and 75 nM d25/r45) with various concentrations of dGTP (0.2, 0.5, 1, 2, and 4 mM). Reactions were stopped at various timepoints within 10 s by addition of 0.5 M EDTA. (B) The same experiments were repeated by hand quench and data were collected within 5 min. (C) The same experiments were repeated by hand quench and data were collected within 30 min. The three experiments performed within different timescales were fit simultaneously according to the one-step binding model shown in Scheme 5.4. Rate constants derived from the global analysis were listed in Table 5.4.

**Rate-determining vs Specificity-determining Step**

In this study, the kinetic pathways of correct and incorrect nucleotide incorporation were identified (Scheme 5.3 and 5.4). Kinetic parameters governing nucleotide binding, incorporation, and concomitant PPi release are listed in Table 5.2 and 5.4. The data showed that PPi release was the rate-determining step during TTP incorporation when an RNA template was used. Previous studies of nucleotide selectivity by HIVRT and T7 DNA polymerase identified an important kinetic partitioning between chemistry ($k_3$) and the reverse of the conformational change ($k_{-2}$) that allows simplifications to be made when deriving the specificity constant (*86-87*). Notably, when $k_3$ is much greater than $k_{-2}$, the specificity constant ($k_{cat}/K_m$) can be simplified to $K_1k_2$ because the closed ternary complex is committed to catalysis. Herein, the kinetic partitioning of $FD_{25}R_{45}$.N state also favors chemistry over nucleotide release ($k_3 >> k_{-2}$). However, when chemistry proceeds, a new kinetic partitioning between PPi release ($k_4$) and the reverse of chemistry ($k_{-3}$) also becomes important. Because $k_4 >> k_{-3}$, the specificity constant governing the incorporation of a normal nucleotide (TTP) is not affected by the slow PPi release; rather it is solely determined by the net rate of binding ($K_1k_2$). Again, the data demonstrate that nucleotide selectivity by HIVRT is determined by the first largely irreversible step, which is the substrate-induced isomerization. The free-energy profile for TTP binding and incorporation then illustrates this analysis. In Figure 5.8A, the nucleotide-induced conformational change ($ED_{25}R_{45}$.N to $FD_{25}R_{45}$.N) represents the highest energy barrier relative to the unbound state ($ED_{25}R_{45}$) and therefore determines the enzyme specificity constant ($k_{cat}/K_m$). Although PPi release reflects the

highest absolute barrier (relative to the local minimum) and therefore represents the rate-limiting step ($k_4$ approximates $k_{cat}$), it does not contribute to enzyme specificity, which is further supported by the fact that $k_{pol}/K_{d,app} \approx k_{cat}/K_m$ (Table 5.2).

The scenario changes during noncognate nucleotide incorporation, in which the product release rate ($k_3$) was shown to reduce $k_{cat}/K_m$, defining the nucleotide specificity constant. The slow PPi release rate allows for the reversal of chemisty ($k_{-2}$), which effectively reduces $k_{cat}/K_m$ by 24-fold according to the constants defined for dGTP incorporation (Table 5.4). The data demonstrated that when $k_{-2} >> k_3$, an equilibrium is established prior to product release, which results in a simplified derivation of specificity corresponding to product of the equilibrium constants and the rate of product release ($K_1K_2k_3$) (*87*). This is further illustrated in the free-energy profile for dGTP incorporation (Figure 5.8B) where the product release reflects the highest barrier relative to the unbound state and the local minimum. Therefore, slow product release represents the rate- and specificity-determining step during incorrect nucleotide incorporation.

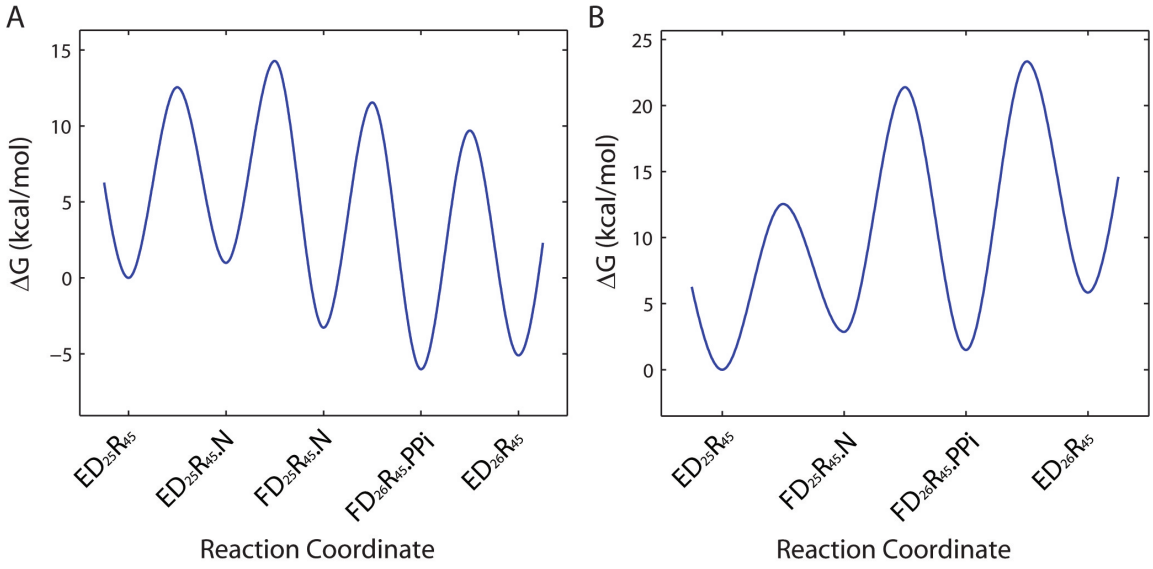



**Figure 5.8** *Free-energy Profiles for TTP (A) and dGTP (B) Incorporation.* (A) The free energy was calculated as $\Delta G = RT[\ln(kT/h) - \ln(k_{obs})]$ kcal/mol using rate constants from Table 5.2. The constant $k$ is the Boltzmann constant, $T$ is 310 K, $h$ is Planck's constant and $k_{obs}$ is the first-order rate constant. Concentrations of nucleotide and PPi were set as 100 and 150 μM, respectively. (B) The free-energy profile for dGTP mis-incorporation was generated using rate constants from Table 5.4.

## 5.4 DISCUSSION

One of the most important results of this report is the identification of rate-limiting PPi release during TTP incorporation by HIVRT with a DNA/RNA hybrid. PPi release has usually been assumed to be fast, so chemistry represents the only rate-limiting step for primer extension. As shown in Figure 1B, the rate of PPi release after incorporation of TTP using a DNA template must be quite fast (> 500 $s^{-1}$) since the fluorescence signal corresponding to PPi is coincident with the observed chemistry. Examination of the kinetics of processive synthesis using a DNA template reveals that each incorporation reaction proceeds at a rate approximately equal to the rate of

chemistry observed in a single turnover (*20, 88*). Therefore, there is no significant delay after the incorporation of one nucleotide and preceding the binding of the next. The inability to show a delay in binding the next nucleotide after a single incorporation event (Figure 1C-F) also supports that translocation must be faster than the rate of incorporation measured in a single turnover.

In contrast, the data revealed that PPi product release represents the rate-limiting step when an RNA template is used for primer extension (Figure 5.2). This conclusion is further supported by the observation of significantly delayed binding of the next nucleotide after the first turnover (Figure 5.4). A rate-limiting translocation seems to be infeasible although the possibility cannot be fully excluded. Moreover, direct measurements of PPi release clearly demonstrate that product release is indeed rate-limiting. Since actual PPi release should be a fast process, it is reasonable to suppose that product release is limited by the isomerization of HIVRT to the open state. This hypothesis was supported by the fact that reopening of RT measured by stopped-flow assays (Figure 5.2B) is coincident with PPi release measured by our coupled stopped-flow assay (Figure 5.2D). Although these data do not demonstrate the order of events, it is reasonable to propose that enzyme opening precedes and therefore limits the rate of PPi release.

Although slow PPi release has been reported during the incorporation of modified or mismatched nucleotides (*89-91*) by the human mitochondrial DNA polymerase, this is the first instance where rate-limiting PPi release is observed using cognate nucleotides. Because of the relatively slow PPi release, the reverse rate constant for chemistry ($k_{-3}$)

was measured and a new kinetic partitioning between product release and the reverse chemistry was identified. Similarly, this partitioning strongly favors the forward product release ($k_4 >> k_{-3}$) and therefore does not contribute to nucleotide selectivity. In chapter 4, our data also showed that the rate-limiting PPi release plays an important role to coordinate the polymerase and RNase H active sites of HIVRT.

The most remarkable observation reported here is the exceedingly slow release of PPi after incorporation of an incorrect nucleotide. The effect of slow PPi release on the efficiency of dGTP incorporation by HIVRT is quite large. The commitment to forward catalysis is not made until PPi dissociates from the enzyme, indicating that the specificity constant for incorporation should be correspondingly low. An unambiguous measurement of $k_{cat}/K_m$ found this to be the case and dictates that previous calculation of the discrimination ($(k_{cat}/K_m)_{TTP}/(k_{cat}/K_m)_{dGTP}$) against dGTP must be revised to $(2.4\pm0.7)\times10^6$, more than 20-fold greater than previously estimated. Similarly, in the case of AZT and 8-oxodG incorporation by human mitochondrial DNA polymerase, slow PPi release allows for the direct reversal of the chemistry step and reduces $k_{cat}$ for incorporation (*89-90*). This represents a novel mechanism by which a polymerase can increase discrimination against an undesirable substrate.

The 20-fold increase of discrimination that we observed for HIVRT may also fill the gap between the fidelity of HIV-1 RT measured by *in vitro* and *in vivo* approaches. More combinations of mis-incorporation will need to be tested in the future to obtain a full picture concerning the effect of slow PPi release on nucleotide selectivity. Similarly, slow PPi release was also observed during incorporation of incorrect nucleotides when a

DNA template was used. Further studies will be performed to evaluate the extent to which the slow PPi release would increase the fidelity of HIV-1 RT during plus strand DNA synthesis since unequal error rates were observed with RNA and DNA templates (*52*).

To get the best estimates for the kinetic parameters governing the incorporation of dGTP, the data obtained over different timescales were fit simultaneously (Figure 5.7), according to the model shown in Scheme 5.4. The simple model adequately accounts for the observed dependence of the amplitude of the fast phase on dGTP concentration. Similar concentration dependence has been observed during AZT and 8-oxodG incorporation by human mitochondrial DNA polymerase (*89-90*). The characteristic pattern suggests that the binding of dGTP is reversibly linked to product formation, a conclusion that requires that PPi release is slow or readily reversible. To explore these possibilities, we first examined the kinetics of the reverse reaction using a dGMP-terminated primer strand. To our surprise, no pyrophosphorolysis reaction was observed. The data revealed that RT was able to catalyze the reverse reaction only when a naturally terminated primer strand was used (Figure 5.2E). The conformational changes of RT during pyrophosphorolysis were not measured, but the fact that we were not able to force the reverse reaction to take place by adding relatively high concentration of PPi to the RT-DNA/RNA complex (primer terminated with dGMP) suggests that when dGMP is in a mismatch with the template strand, PPi is unable to productively associate with the RT-DNA/RNA complex with sufficient binding energy to organize the active site to promote pyrophosphorolysis. These results argued against the hypothesis that a substantial reverse

rate constant for phosphate binding and reaction was the cause of the abnormal kinetics of dGTP incorporation. Rather the data suggest that slow PPi release is limited by the enzyme opening following incorporation of a mismatch, which provides an additional checkpoint for improving enzyme fidelity. The slow PPi release provide additional time for the reversal of chemistry and release of the bound mismatched nucleotide.

Currently, we have no direct measurement to demonstrate that the slow phase in Figure 5.7B corresponds to the PPi release step. The coupled stopped-flow assay was performed to measure the rate of PPi dissociation using dGTP as the substrate. However, a considerable overestimation of the rate was likely obtained from the data because the presence of the phosphate mop would scavenge traces of phosphate with a faster rate than its release from the reaction and therefore result in reduced amplitude (*181*). This explains the failure to get a good fit when including the fluorescence data. In the future, a different experimental approach will be used to directly measure the rate of PPi release by synthesizing $\gamma$-$^{32}$P-labeled dGTP (*90*).

In summary, here we show for the first time that PPi release limits the rate of processive RNA-dependent DNA synthesis and improves the overall fidelity of the enzyme. These findings enriched our understandings of the kinetic basis for the rate-determining versus specificity-determining step and allow estimates for the rate of reversal of the chemistry step at the active site of the enzyme. Accordingly, these data provide the first measurement of the equilibrium constants governing polymerization and the first complete free energy profile for HIVRT.

# Chapter 6: Kinetic Characterization of HIV Reverse Transcriptase R72A Mutant

## 6.1 INTRODUCTION

HIVRT is an RNA- and DNA-dependent DNA polymerase with moderate fidelity. Given the lack of proofreading activity, nucleotide selectivity is solely a function of the binding and reaction of nucleotide at the polymerase active site. Although geometric constraints of the nucleotide binding pocket may play an important role in discriminating against non-complementary base pairs (*35*), numerous biochemical and structural studies have been performed to determine the amino acid residues involved in the binding of the incoming nucleotide (*35, 182-185*). The pocket includes the conserved YMDD motif (Y183, M184, D185, and D186), three basic residues that are in the vicinity of the phosphates of dNTP (K65, R72, and K219), and the 'steric gate' Y115. Other important residues include D113 and Q151. Among these residues, D185 and D186 were well studied since they have important catalytic functions (*185*). Residues associated with NRTIs resistance also have been extensively studied, e.g., K65 (*70, 186*), Q151 (*187-189*), M184 (*190-194*), and K219 (*108, 195*). Despite its functional importance, kinetic studies of R72 were only sparse. One study showed that a R72A mutation strongly altered the fidelity of HIVRT (*196*), the kinetic basis for which was unfortunately missing. Another study showed that R72 might be involve in the pyrophosphate binding/release function of HIV RT (*37*), probably due to the hydrogen bonding between its guanidinium group and the $\alpha$-phosphate of the incoming nucleotide (*35*). However, this conclusion

was based on indirect evidence so a direct and accurate measurement of the rate of pyrophosphate release is of great importance to test the conclusion.

In this chapter, pre-steady-state kinetic assays and global data analysis were performed to determine how the R72A mutation would affect the kinetics of each elementary step during nucleotide incorporation. Then, incorporation kinetics of correct and incorrect nucleotides were compared to explore the mechanistic basis for the altered fidelity by HIVRT_R72A mutant.

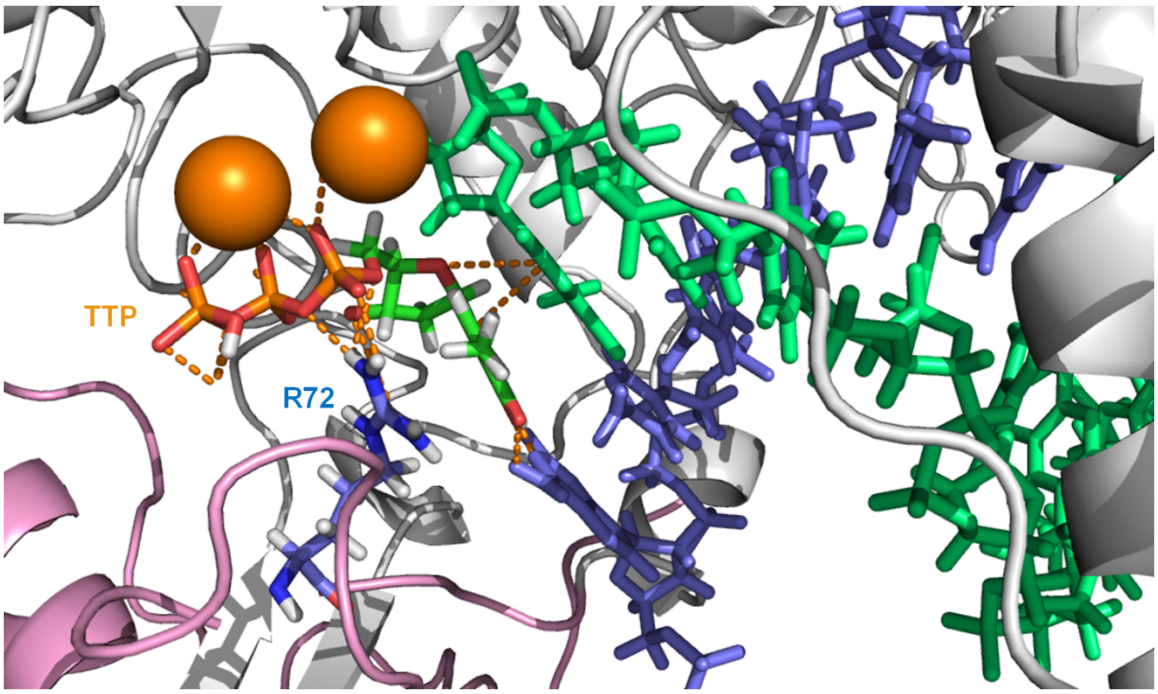



**Figure 6.1** *Hydrogen bonding between residue R72 and the α-phosphate of TTP*. The crystal structure of HIVRT (PDB: 1RTD) in complex with a DNA primer/template (green and blue) and TTP (denoted by orange and color coded by element). Binding of TTP induces the closure of the finger domain (pink). The TTP base stacks against the base at the 3'-end of the primer while the triphosphate wraps around one of the $Mg^{2+}$ ion (orange). Residue R72 (denoted by blue and CPK color coded by element) interacts with the α-phosphate through hydrogen bonding.

## 6.2 MATERIALS AND METHODS

### Expression and Purification of Proteins Used in This Study

HIVRT R72A mutant was expressed, purified and labeled following methods previously described (*77, 86*). Briefly, single mutations were introduced in genes of both subunits using a Quikchange Multi kit (Stratagene). The method for the preparation of MDCC-labeled PBP used in this study was described previously (*154*) and its active site concentration was determined by fluorescent titration with various concentrations of inorganic Pi.

### Pre-steady-state Kinetic Measurements

Single-nucleotide incorporation was monitored by rapid quench flow and stopped flow methods. Chemical quench flow incorporation experiments were performed to determine the incorporation efficiency of TTP by R72A mutant. Various concentrations of TTP (10, 25, 50, 100, and 200 μM) were rapidly mixed with a preformed enzyme-DNA/RNA complex (175 nM MDCC-labeled HIVRT_R72A and 75 nM d25/r45 (Table 6.1)) using a RQF-3 Rapid Quench Flow (KinTek Corp.). All reactions were conducted in 50 mM Tris pH7.5, 100 mM KCl and 10 mM Magnesium acetate (all concentrations final) at 37°C and stopped by adding 0.5 M EDTA. Products were separated and quantified as described above.

The time dependence of transient MDCC fluorescence changes corresponding to TTP induced enzyme conformational change was monitored using an Auto SF-200x (KinTek Corp.). A preformed enzyme-DNA/RNA duplex (100 nM of MDCC-labeled HIVRT_R72A and 150 nM $d25_{ddA}$/r45 (Table 6.1)) was rapidly mixed with various

concentrations of TTP (20, 30, 40, 60, 80, 100, 150, and 200 μM) in the stopped flow for 300 ms. MDCC fluorescence was excited at 425 nm and fluorescence was monitored at 475 nm with a 50 nm band-pass filter (Semrock). To monitor the fluorescence changes corresponding to the chemistry step, a normal primer (not dideoxy-terminated) was used to form the enzyme-DNA/RNA binary complex (100 nM of MDCC-labeled HIVRT_R72A and 150 nM d25/r45 (Table 6.1)) and mixed with various concentrations of TTP (5, 10, 20, 30, 40, 60, 80, 100 and 150 μM) in the stopped flow for a longer time scale of 50 s.

A coupled fluorescence assay described elsewhere (*125*) was used to measure the kinetics of pyrophosphate (PPi) release. Briefly, the preformed enzyme-DNA/RNA duplex (150 nM of HIVRT mutant and 100 nM d25/r45) was rapidly mixed with 25 μM of TTP in the presence of 0.6 μM of pyrophosphatase (PPase), the Pi "mop" containing 100 μM of 7-methylguanosine (7-MEG) and 0.02 IU/mL purine nucleoside phosphorylase (PNPase), as well as 1.5 μM of fluorescently labeled E. coli PBP.

**Table 6.1** *DNA/RNA Hybrids for Kinetic Assays*

*$d25_{ddA}$/r45:*

$d25_{dd}$: 5′-GCCTCCCCCCCCCCCAACCAACTCA$_{dd}$-3′
r45 : 3′-CGGAGGGGGGGGGGGUUGGUUGAGU**A**GCAGCUAGGUUACGGCAGG-5′

*d25/r45:*

d25: 5′-GCCTCCCCCCCCCCCAACCAACTCA-3′
r45: 3′-CGGAGGGGGGGGGGGUUGGUUGAGU**A**GCAGCUAGGUUACGGCAGG-5′

**Quench Flow dGTP:dA Incorporation**

The incorporation efficiency ($k_{pol}/K_{d,app}$) of dGTP by HIVRT_R72A was also determined by performing quench flow incorporation assays. A preformed enzyme-DNA/RNA complex (175 nM MDCC-labeled HIVRT_R72A and 75 nM d25/r45 (Table 6.1)) was rapidly mixed with various concentrations of dGTP (0.5, 1, 2, 4, and 8 mM) using a RQF-3 Rapid Quench Flow (Kin Tek Corp.). Reactions were stopped by adding 0.5 M EDTA at various time points up to 20 min. To evaluate the effect of R72A mutation on the fidelity of HIVRT, specificity constants ($k_{pol}/K_{d,app}$) of TTP and dGTP incorporation by HIVRT_WT protein were also determined through quench flow assays. For TTP incorporation, a preformed enzyme-DNA/RNA complex (175 nM HIVRT_WT and 75 nM d25/r45 (Table 6.1)) was rapidly mixed with various concentrations of TTP (0.5, 1.5, 6, 20, and 60 μM). Reaction times ranged from 0 to 1 s. As to dGTP incorporation, a preformed enzyme-DNA/RNA complex (175 nM HIVRT_WT and 75 nM d25/r45 (Table 6.1)) was rapidly mixed with various concentrations of dGTP (0.1, 0.2, 0.5, 1, and 2 mM). Reaction times ranged from 0 to10 s.

**Data Analysis**

Quench flow incorporation, stopped flow binding and incorporation, and stopped flow PPi release data obtained with HIVRT_R72A were fit globally to a four-step kinetic model using KinTek Explorer software to Scheme 6.1. The PPi release step was assumed to be irreversible since the data set was insufficient to define the small rate of PPi rebinding. Nucleotide incorporation data were analyzed by fitting to the minimum model that includes a one-step, rapid-equilibrium nucleotide binding event followed by the rate-limiting chemistry step (Scheme 6.2). Specificity constant ($k_{pol}/K_{d,app}$) values were

derived from the simulations. Discrimination is defined by D = $(k_{pol}/K_{d,app})_{TTP}$ / $(k_{pol}/K_{d,app})_{dGTP}$. The influence of enzyme specificity by introducing the R72A mutation is then defined by $I = D_{R72A}/D_{WT}$.

**Scheme 6.1** *Two-step Nucleotide Binding, Incorporation, and PPi Release*

$$ED_{25} + N \underset{}{\overset{1/K_1}{\rightleftharpoons}} ED_{25}N \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} FD_{25}N \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} FD_{26}{\cdot}PP \xrightarrow{k_4} ED_{26} + PP$$

**Scheme 6.2** *Minimum Model of Nucleotide Incorporation*

$$ED_{25} + N \overset{K_{d,app}}{\rightleftharpoons} ED_{25}N \overset{k_{pol}}{\rightleftharpoons} ED_{26} + PP$$

## 6.3 RESULTS

### Kinetic Characterization of HIVRT_R72A Mutant

Stopped flow assays were first performed to measure the kinetics of a single turnover of TTP incorporation (Figure 6.2A). Various concentrations of TTP (5, 10, 20, 30, 40, 60, 80, 100 and 150 μM) were rapidly mixed with a preformed enzyme-DNA/RNA complex (100 nM MDCC-labeled HIVRT_R72A and 150 nM d25/r45). In contrast to what we observed before, only a decrease of fluorescence was recorded, with a half-life around 6 s at the highest concentration of TTP. The slow drop of fluorescent signal was not consistent with the nucleotide induced fast closure of HIVRT fingers domain. Consequently, TTP-induced conformational change was measured with a dideoxy-terminated primer to block the chemistry step (Figure 6.2B). Various concentrations of TTP (20, 30, 40, 60, 80, 100, 150, and 200 μM) were mixed with a preformed enzyme-$DNA_{dd}$/RNA complex (100 nM MDCC-labeled HIVRT_R72A and

150 nM d25ddA/r45). Our data showed an increase of fluorescence signal upon the binding of TTP, indicating that the drop of fluorescence observed in panel A was probably due to the chemistry step. Next, manual quench flow assays were performed to measure the apparent binding affinity of TTP ($K_{d,app}$) and the rate of polymerization ($k_{pol}$) by HIVRT_R72A mutant (Figure 6.2C). As expected, the chemistry step measured here was as slow as the fluorescence decrease in panel A, further supporting that the unusual decrease of fluorescence was due to the chemistry step. Finally, a coupled assay was performed to measure the kinetics of PPi release (Figure 6.2D) since indirect evidence has shown that the R72A mutation significantly affected the PPi release step (*37*). Global data analysis of all four experiments was performed to derive rate constants governing TTP binding and incorporation. The TTP ground-state binding constant ($1/K_1$) was determined as 92 ± 3 μM, which was 3-fold tighter than its binding to a WT protein (*77*). The forward ($k_2$) and reverse ($k_{-2}$) rates for the conformational change were 127 ± 2 and 32 ± 1 $s^{-1}$, which defined a $K_2$ value of 4 ± 0.2 which was 170-fold smaller than that of TTP incorporation by WT_HIVRT-MDCC (690±67) (*77*), suggesting that R72 might play a crucial role in the nucleotide induced isomerization step. Rates of chemistry and PPi release were determined as 0.18 ± 0.003 and 0.04 ± 0.008 $s^{-1}$, respectively. The rate of chemistry was reduced by 190-fold due to the R72A substitution. The rate of PPi release during TTP incorporation by WT_HIVRT protein was too fast to be accurately determined and a lower limit of 500 $s^{-1}$ was obtained from the confidence contour analysis (Chapter 5). However, the R72A mutation dramatically impaired PPi release, making its rate 4-fold slower than that of chemistry.

The net binding affinity of TTP ($K_{d,net}$) as a result of two-step nucleotide binding was calculated as $1/(K_1(1+K_2) = 18 \pm 1$ μM, which was 45-fold weaker than its binding to a WT_HIVRT protein. The reverse conformational change rate ($32 \pm 1$ $s^{-1}$) is greater than the rate of chemistry ($0.18 \pm 0.003$ $s^{-1}$), which allows the two-step binding mechanism come to equilibrium, leading to a reduced $K_{d,app}$ value ($21 \pm 0.3$ μM) that resembles the net $K_d$. The rate of polymerization ($k_{pol}$) is $0.12 \pm 0.001$ $s^{-1}$, yielding a $k_{pol}/K_{d,app}$ value as $(6 \pm 0.1)\times10^{-3}$ $\mu M^{-1}$ $s^{-1}$. Rate of PPi release ($0.04 \pm 0.008$ $s^{-1}$) is comparable to the reverse chemistry step rate ($0.07 \pm 0.002$ $s^{-1}$), suggesting that PPi removal may also contribute to enzyme specificity. The specificity constant ($k_{cat}/K_m$) was calculated as $k_{cat}/K_m = k_1k_2k_3k_4/(k_2k_3k_4+k_{-1}(k_{-2}k_{-3}+k_{-2}k_4+k_3k_4)) = (3 \pm 0.2)\times10^{-3}$ $\mu M^{-1}s^{-1}$, with a $k_{cat}$ value of $0.023 \pm 0.001$ $s^{-1}$ and a $K_m$ value as $8 \pm 0.3$ μM. All kinetic parameters were summarized in Table 6.2.

**Table 6.2** *Kinetic Parameters Governing TTP Binding and Incorporation by HIVRT_R72A Mutant*

| $1/K_1$ ($\mu M$) | $k_2$ ($s^{-1}$) | $k_{-2}$ ($s^{-1}$) | $K_2$ | $k_3$ ($s^{-1}$) | $k_{-3}$ ($s^{-1}$) | $k_4$ ($s^{-1}$) |
|---|---|---|---|---|---|---|
| 92±3 | 127±2 | 32±1 | 4±0.2 | 0.18±0.003 | 0.07±0.002 | 0.04±0.008 |
| $K_{d,net}$[a] ($\mu M$) | $k_{pol}$ ($s^{-1}$) | $K_{d,app}$ ($\mu M$) | $k_{pol}/K_{d,app}$ ($\mu M^{-1}s^{-1}$) | $k_{cat}$ ($s^{-1}$) | $K_m$ ($\mu M$) | $k_{cat}/K_m$[b] ($\mu M^{-1}s^{-1}$) |
| 18±1 | 0.12±0.001 | 21±0.3 | $(6\pm0.1)\times10^{-3}$ | 0.023±0.001 | 8±0.3 | $(3\pm0.2)\times10^{-3}$ |

[a] $K_{d,net} = 1/(K_1(1+K_2))$

[b] $k_{cat}/K_m = k_1k_2k_3k_4/(k_2k_3k_4+k_{-1}(k_{-2}k_{-3}+k_{-2}k_4+k_3k_4))$

**Kinetics of Mismatch Incorporation and Enzyme Specificity**

Kinetics of dGTP:dA misincorporation by R72A mutant were examined by the manual quench flow method (Figure 6.3A). A preformed enzyme-DNA/RNA complex (175 nM MDCC-labeled HIVRT_R72A and 75 nM d25/r45) was mixed with various concentrations of dGTP (0.5, 1, 2 and 4 mM). The concentration-dependence of dGTP incorporation was analyzed by fitting to the minimum model shown in Scheme 6.2. The $K_{d,app}$ was 1.3 ± 0.1 mM, and the $k_{pol}$ was 0.002 ± 0.0004 $s^{-1}$, yielding a $k_{pol}/K_{d,app}$ value as $(1.5 \pm 0.3)\times10^{-6}$ $\mu M^{-1}s^{-1}$. The specificity constants ($k_{pol}/K_{d,app}$) governing TTP and dGTP incorporation by HIVRT_R72A were $(6 \pm 0.1)\times10^{-3}$ and $(1.5 \pm 0.3)\times10^{-6}$ $\mu M^{-1}s^{-1}$, respectively. These values produced a net discrimination $[(k_{pol}/K_{d,app})_{TTP} / (k_{pol}/K_{d,app})_{dGTP}]$ of $(4 \pm 0.8)\times10^3$ (Table 6.3).

The incorporation of TTP (Figure 6.3B) and dGTP (Figure 6.3C) by WT_HIVRT was also examined to evaluate the extent to which the R72A mutation could affect the enzyme specificity. Briefly, a preformed enzyme-DNA/RNA complex (175 nM MDCC-labeled WT_HIVRT and 75 nM d25/r45) was mixed with various concentrations of

nucleotides (TTP: 0.25, 0.75, 3, 10, 30 μM; or dGTP: 0.1, 0.2, 0.5, 1, and 2 mM). The specificity constants ($k_{pol}/K_{d,app}$) governing TTP and dGTP incorporation by WT_HIVRT were 6.5 ± 0.8 and (5.2 ± 1.2)×$10^{-4}$ $\mu M^{-1}s^{-1}$, respectively. These values defined the dGTP discrimination by the WT_HIVRT of (1.3 ± 0.3)×$10^4$. The influence of R72A mutation on enzyme specificity is defined by $I = D_{R72A}/D_{WT} = (4 \pm 0.8)\times10^3 / (1.3 \pm 0.3)\times10^4 = 0.3 \pm 0.1$, indicating that R72A decreased the specificity of HIVRT by ~3-fold (Table 6.3).

**Table 6.3** *Specificity and Discrimination for WT and R72A HIVRT*

| Enzyme | dNTP | $k_{pol}$ ($s^{-1}$) | $K_{d,app}$ (μM) | $k_{pol}/K_{d,app}$ ($\mu M^{-1}s^{-1}$) | D[a] | I[b] |
|---|---|---|---|---|---|---|
| HIVRT_R72A | TTP | 0.12±0.001 | 21±0.3 | (6±0.1)×$10^{-3}$ | | |
| | dGTP | 0.002±0.0004 | 1300±100 | (1.5±0.3)×$10^{-6}$ | (4±0.8)×$10^3$ | |
| HIVRT_WT | TTP | 142±12 | 22±2 | 6.5±0.8 | | |
| | dGTP | 1.3±0.2 | 2500±400 | (5.2±1.2)×$10^{-4}$ | (1.3±0.3)×$10^4$ | 0.3±0.1 |

[a] $D = [(k_{pol}/K_{d,app})_{TTP} / (k_{pol}/K_{d,app})_{dGTP}]$
[b] $I = D_{R72A}/D_{WT}$

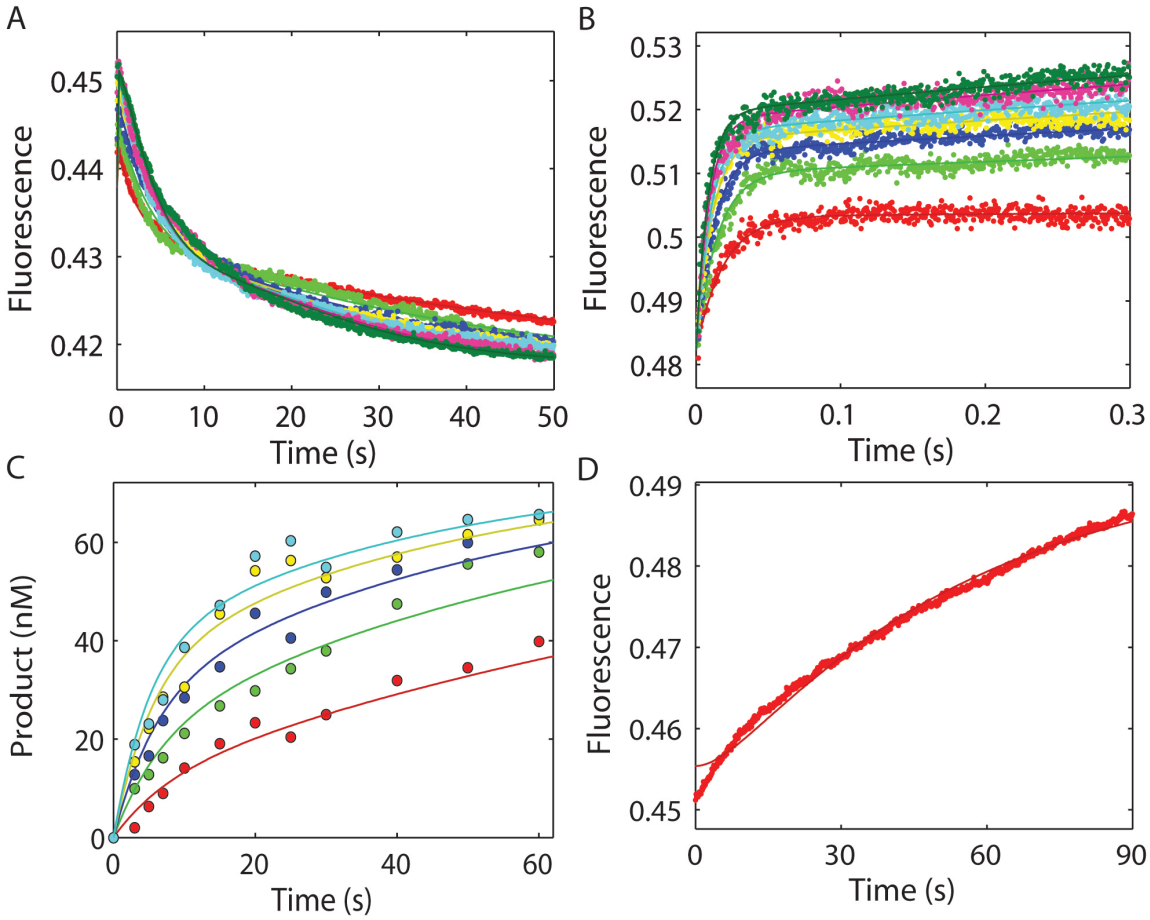

A
Fluorescence
0.45
0.44
0.43
0.42
0
10
20
30
40
50
Time (s)
B
Fluorescence
0.53
0.52
0.51
0.5
0.49
0.48
0
0.1
0.2
0.3
Time (s)
C
Product (nM)
60
40
20
0
0
20
40
60
Time (s)
D
Fluorescence
0.49
0.48
0.47
0.46
0.45
0
30
60
90
Time (s)

**Figure 6.2**

**Figure 6.2** *Kinetic characterization of MDCC-labeled HIVRT R72A mutant by global analysis of TTP binding, incorporation and PPi release*. (A) The time dependence of the fluorescence change during TTP incorporation was monitored after mixing various concentrations of TTP (5, 10, 20, 30, 40, 60, 80, 100 and 150 μM) with a preformed enzyme-DNA/RNA complex (100 nM MDCC-labeled HIV RT_R72A and 150 nM d25/r45). (B) TTP induced conformational change was measured with dideoxy-terminated primer to prohibit the following chemistry step. Various concentrations of TTP (20, 30, 40, 60, 80, 100, 150, and 200 μM) were mixed with a preformed enzyme-$DNA_{dd}$/RNA complex (100 nM MDCC-labeled HIVRT_R72A and 150 nM d25ddA/r45). (C) The concentration-dependence of TTP incorporation was measured using the manual quench flow method. Various concentrations of TTP (10, 25, 50, 100, and 200 μM) were mixed with a preformed enzyme-DNA/RNA complex (175 nM MDCC-labeled HIVRT_R72A and 75 nM d25/r45). (D) Rate of PPi release during TTP incorporation was monitored by rapidly mixing 25 μM of TTP with a preformed enzyme-DNA/RNA complex (150 nM MDCC-labeled HIVRT_R72A and 100 nM d25/r45) in the presence of 1.5 μM of MDCC-PBP and 0.6 μM of PPase. Fluorescence was observed by excitation of MDCC at 425 nm and monitoring emission with a 475 nm band-pass filter with a 50 nm bandwidth. Global data analysis of all four experiments was performed to derive rate constants governing TTP binding and incorporation. The TTP ground-state binding constant ($1/K_1$) was determined as 92 ± 3 μM. The forward ($k_2$) and reverse ($k_{-2}$) rates for the conformational change was 127 ± 2 and 32 ± 1 $s^{-1}$, which defined a $K_2$ value of 4 ± 0.2 that was 170-fold smaller than that of WT HIVRT-MDCC (690 ± 67) (*77*). Rates of chemistry and PPi release were determined as 0.18 ± 0.003 and 0.04 ± 0.008 $s^{-1}$, respectively.

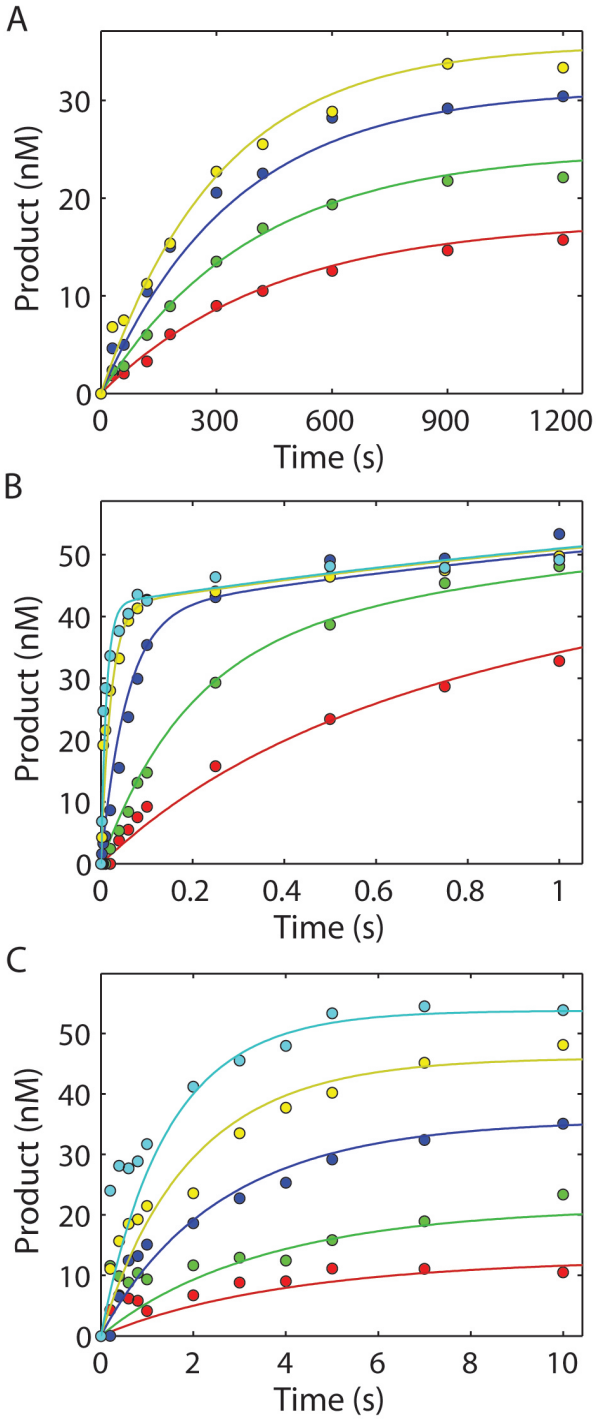



**Figure 6.3**. *Specificity and Discrimination for WT and R72A HIVRT*. (A) Kinetics of dGTP:dA misincorporation by R72A mutant were examined by the manual quench flow method. A preformed enzyme-DNA/RNA complex (175 nM MDCC-labeled HIVRT_R72A and 75 nM d25/r45) was mixed with various concentrations of dGTP (0.5, 1, 2 and 4 mM). The concentration-dependence of dGTP incorporation was analyzed by fitting to the minimum model shown in Scheme 6.2. The $K_{d,app}$ was 1.3 ± 0.1 mM, and the $k_{pol}$ was 0.002 ± 0.0004 $s^{-1}$, yielding a $k_{pol}/K_{d,app}$ value as (1.5 ± 0.3)×$10^{-6}$ $\mu M^{-1}s^{-1}$. The incorporation of TTP (B) and dGTP (C) by WT_HIVRT was also examined to evaluate the extent to which the R72A mutation could affect the enzyme specificity. Briefly, a preformed enzyme-DNA/RNA complex (175 nM MDCC-labeled WT_HIVRT and 75 nM d25/r45) was mixed with various concentrations of TTP (0.25, 0.75, 3, 10, 30 μM) or dGTP (0.1, 0.2, 0.5, 1, and 2 mM). The specificity constants ($k_{pol}/K_{d,app}$) governing TTP and dGTP incorporation by WT_HIVRT were 6.5 ± 0.8 and (5.2 ± 1.2)×$10^{-4}$ $\mu M^{-1}s^{-1}$, respectively. The kinetic parameters were summarized in Table 6.3.

## 6.4 DISCUSSION

In this chapter, the HIVRT_R72A mutant was kinetically characterized, which revealed that this mutation affected nucleotide binding, conformational change, chemistry, and PPi release to different extents. Since nucleotide-induced conformational changes played an important role in enzyme specificity, this mutation also resulted in decreased fidelity.

First, several quench flow and stopped flow assays were performed, together with global data analysis, to define the kinetic pathway of TTP incorporation by R72A mutant. In the stopped flow assays to measure kinetics of TTP binding and incorporation, an unusual pattern of fluorescence change was observed: the fluorescence increased when the binding of nucleotide induced fingers closure while decreased when the ternary complex reopened after chemistry (Figure 6.2A-B). This assignment was further supported by the fact that the fluorescence signal decreased with the same rate as the catalytic step (Figure 6.2A and C). This unusual pattern was previously observed during dATP:dA mis-incorporation by HIVRT (*58*), and MD simulations in the same study predicted a disordered, higher free energy state distinct from the closed state induced by the binding of a correct nucleotide. An less favorable conformational change step ($K_2$ = 4 ± 0.2) was also observed in this study during TTP incorporation by R72A mutant, suggesting that the fingers domain of this mutant might adopt a similar conformation that disfavors catalysis and promotes release. Structural studies of DNA polymerases reveal that the binding of an incorrect nucleotide can induce an intermediate conformation that is distinct from the established 'open' and 'closed' states (*197-198*). This 'partially closed' state might represent the fingers conformation of the R72A mutant after the binding of TTP and result in the high fluorescent state.

Previous studies have shown that the R72A substitution impaired both catalytic and PPi release steps during nucleotide incorporation (*37*). While data presented in this chapter demonstrated that the mutation affected all steps in the kinetic pathway for TTP incorporation. The ground state binding of TTP was 3-fold tighter whereas the isomerization step was 170-fold less favorable so that the overall net binding was 45-fold weaker. Structural details for the impaired nucleotide-induced isomerization step are still unknown. However, a nearby residue W71 also plays an important role in this step, suggesting that this region is crucial for the induced-fit mechanism and therefore remains conserved during evolution. As expected, the rate of chemistry was reduced by 190-fold, while PPi release was even slower and became the rate-limiting step.

In the field of DNA polymerase, $k_{pol}/K_{d,app}$ (instead of $k_{cat}/K_m$) has been widely accepted as the specificity constant, which is valid when nucleotide binding is a one-step, rapid equilibrium process and when chemistry is the only rate-limiting step for product formation. However, this substitution is no longer appropriate when PPi release becomes the rate-limiting step. For instance, the $k_{pol}/K_{d,app}$ value is 2-fold larger than $k_{cat}/K_m$ due to the rate-limiting PPi release step (Table 6.2). This issue is detailed in chapter 5 and is demonstrated again here.

Data presented in this chapter revealed that R72A mutation reduced the fidelity of HIVRT by ~3-fold, which is consistent with the important role of induce-fit in enzyme specificity. However, a previous study proposed that R72A substitution increased the fidelity and the mutant was a base substitution antimutator (*196*). This increased nucleotide selectivity, however, was sequence dependent, which might explain the reduced fidelity observed here. Also, previous data have shown that slow PPi release provided extra fidelity (Chapter 5). Unfortunately, this step was too slow to be accurately measured during mis-incorporation (due to the presence of Pi mop), especially in the case

of R72A mutation that slowed the PPi release step further. In the future, a new assay will be developed to measure this extremely slow PPi release step by attaching a $^{32}P$ label on the γ-phosphate of the incoming nucleotide so that kinetics of PPi release can be resolved by quench flow assays that are insensitive to Pi contaminations.

## References


1. Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., and Montagnier, L. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS), *Science 220*, 868-871.
2. Gallo, R. C., Sarin, P. S., Gelmann, E. P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V. S., Mann, D., Sidhu, G. D., Stahl, R. E., Zolla-Pazner, S., Leibowitch, J., and Popovic, M. (1983) Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS), *Science 220*, 865-867.
3. WHO. (2015) Global summary of the AIDS epidemic - 2014 URL: http://www.who.int/hiv/data/epi_core_july2015.png?ua=1.
4. Lucas, S., and Nelson, A. M. (2015) HIV and the spectrum of human disease, *J Pathol 235*, 229-241.
5. WHO. (2015) Global health sector response to HIV, 2000-2015: focus on innovations in Africa URL: http://apps.who.int/iris/bitstream/10665/198065/1/9789241509824_eng.pdf?ua=1.
6. Pepin, J. (2011) *The origins of AIDS*, Cambridge University Press.
7. Chinen, J., and Shearer, W. T. (2002) Molecular virology and immunology of HIV infection, *J Allergy Clin Immunol 110*, 189-198.
8. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., and et al. (1985) Complete nucleotide sequence of the AIDS virus, HTLV-III, *Nature 313*, 277-284.
9. Fanales-Belasio, E., Raimondo, M., Suligoi, B., and Butto, S. (2010) HIV virology and pathogenetic mechanisms of infection: a brief overview, *Ann Ist Super Sanita 46*, 5-14.
10. Emerman, M., and Malim, M. H. (1998) HIV-1 regulatory/accessory genes: keys to unraveling viral and host cell biology, *Science 280*, 1880-1884.
11. Freed, E. O. (1998) HIV-1 gag proteins: diverse functions in the virus life cycle, *Virology 251*, 1-15.
12. Freed, E. O. (2001) HIV-1 replication, *Somat Cell Mol Genet 26*, 13-33.
13. Sierra, S., Kupfer, B., and Kaiser, R. (2005) Basics of the virology of HIV-1 and its replication, *J Clin Virol 34*, 233-244.
14. Hu, W. S., and Hughes, S. H. (2012) HIV-1 reverse transcription, *Cold Spring Harb Perspect Med 2*.
15. Sarafianos, S. G., Das, K., Tantillo, C., Clark, A. D., Jr., Ding, J., Whitcomb, J. M., Boyer, P. L., Hughes, S. H., and Arnold, E. (2001) Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA, *EMBO J 20*, 1449-1461.
16. Esposito, F., Corona, A., and Tramontano, E. (2012) HIV-1 Reverse Transcriptase Still Remains a New Drug Target: Structure, Function, Classical Inhibitors, and

New Inhibitors with Innovative Mechanisms of Actions, *Mol Biol Int 2012*, 586401.
17. Sarafianos, S. G., Marchand, B., Das, K., Himmel, D. M., Parniak, M. A., Hughes, S. H., and Arnold, E. (2009) Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition, *J Mol Biol 385*, 693-713.
18. Schultz, S. J., and Champoux, J. J. (2008) RNase H activity: structure, specificity, and function in reverse transcription, *Virus Res 134*, 86-103.
19. DeStefano, J. J., Buiser, R. G., Mallaber, L. M., Myers, T. W., Bambara, R. A., and Fay, P. J. (1991) Polymerization and RNase H activities of the reverse transcriptases from avian myeloblastosis, human immunodeficiency, and Moloney murine leukemia viruses are functionally uncoupled, *J Biol Chem 266*, 7423-7431.
20. Kati, W. M., Johnson, K. A., Jerva, L. F., and Anderson, K. S. (1992) Mechanism and fidelity of HIV reverse transcriptase, *J Biol Chem 267*, 25988-25997.
21. DeStefano, J. J., Mallaber, L. M., Fay, P. J., and Bambara, R. A. (1994) Quantitative analysis of RNA cleavage during RNA-directed DNA synthesis by human immunodeficiency and avian myeloblastosis virus reverse transcriptases, *Nucleic Acids Res 22*, 3793-3800.
22. Gotte, M., Fackler, S., Hermann, T., Perola, E., Cellai, L., Gross, H. J., Le Grice, S. F., and Heumann, H. (1995) HIV-1 reverse transcriptase-associated RNase H cleaves RNA/RNA in arrested complexes: implications for the mechanism by which RNase H discriminates between RNA/RNA and RNA/DNA, *EMBO J 14*, 833-841.
23. Huber, H. E., and Richardson, C. C. (1990) Processing of the primer for plus strand DNA synthesis by human immunodeficiency virus 1 reverse transcriptase, *J Biol Chem 265*, 10565-10573.
24. Luo, G. X., Sharmeen, L., and Taylor, J. (1990) Specificities involved in the initiation of retroviral plus-strand DNA, *J Virol 64*, 592-597.
25. Rezende, L. F., and Prasad, V. R. (2004) Nucleoside-analog resistance mutations in HIV-1 reverse transcriptase and their influence on polymerase fidelity and viral mutation rates, *Int J Biochem Cell Biol 36*, 1716-1734.
26. Svarovskaia, E. S., Cheslock, S. R., Zhang, W. H., Hu, W. S., and Pathak, V. K. (2003) Retroviral mutation rates and reverse transcriptase fidelity, *Front Biosci 8*, d117-134.
27. Mansky, L. M. (1996) Forward mutation rate of human immunodeficiency virus type 1 in a T lymphoid cell line, *AIDS Res Hum Retroviruses 12*, 307-314.
28. Mansky, L. M., and Temin, H. M. (1995) Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase, *J Virol 69*, 5087-5094.
29. Abram, M. E., Ferris, A. L., Shao, W., Alvord, W. G., and Hughes, S. H. (2010) Nature, position, and frequency of mutations made in a single cycle of HIV-1 replication, *J Virol 84*, 9864-9878.

30. di Marzo Veronese, F., Copeland, T. D., DeVico, A. L., Rahman, R., Oroszlan, S., Gallo, R. C., and Sarngadharan, M. G. (1986) Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV, *Science 231*, 1289-1291.
31. Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992) Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor, *Science 256*, 1783-1790.
32. Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Jr., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., and et al. (1993) Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA, *Proc Natl Acad Sci U S A 90*, 6320-6324.
33. Zheng, X., Perera, L., Mueller, G. A., DeRose, E. F., and London, R. E. (2015) Asymmetric conformational maturation of HIV-1 reverse transcriptase, *Elife 4*.
34. Zheng, X., Pedersen, L. C., Gabel, S. A., Mueller, G. A., DeRose, E. F., and London, R. E. (2016) Unfolding the HIV-1 reverse transcriptase RNase H domain - how to lose a molecular tug-of-war, *Nucleic Acids Res 44*, 1776-1788.
35. Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance, *Science 282*, 1669-1675.
36. Singh, K., Marchand, B., Kirby, K. A., Michailidis, E., and Sarafianos, S. G. (2010) Structural Aspects of Drug Resistance and Inhibition of HIV-1 Reverse Transcriptase, *Viruses 2*, 606-638.
37. Sarafianos, S. G., Pandey, V. N., Kaushik, N., and Modak, M. J. (1995) Site-directed mutagenesis of arginine 72 of HIV-1 reverse transcriptase. Catalytic role and inhibitor sensitivity, *J Biol Chem 270*, 19729-19735.
38. Steitz, T. A., and Steitz, J. A. (1993) A general two-metal-ion mechanism for catalytic RNA, *Proc Natl Acad Sci U S A 90*, 6498-6502.
39. Nakamura, T., Zhao, Y., Yamagata, Y., Hua, Y. J., and Yang, W. (2012) Watching DNA polymerase eta make a phosphodiester bond, *Nature 487*, 196-201.
40. Vyas, R., Reed, A. J., Tokarsky, E. J., and Suo, Z. (2015) Viewing Human DNA Polymerase beta Faithfully and Unfaithfully Bypass an Oxidative Lesion by Time-Dependent Crystallography, *J Am Chem Soc 137*, 5225-5230.
41. Freudenthal, B. D., Beard, W. A., Shock, D. D., and Wilson, S. H. (2013) Observing a DNA polymerase choose right from wrong, *Cell 154*, 157-168.
42. Cowan, J. A., Ohyama, T., Howard, K., Rausch, J. W., Cowan, S. M., and Le Grice, S. F. (2000) Metal-ion stoichiometry of the HIV-1 RT ribonuclease H domain: evidence for two mutually exclusive sites leads to new mechanistic insights on metal-mediated hydrolysis in nucleic acid biochemistry, *J Biol Inorg Chem 5*, 67-74.

43. Nowotny, M., Gaidamakov, S. A., Crouch, R. J., and Yang, W. (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis, *Cell 121*, 1005-1016.
44. Nowotny, M., and Yang, W. (2006) Stepwise analyses of metal ions in RNase H catalysis from substrate destabilization to product release, *EMBO J 25*, 1924-1933.
45. Yang, W., Lee, J. Y., and Nowotny, M. (2006) Making and breaking nucleic acids: two-Mg2+-ion catalysis and substrate specificity, *Mol Cell 22*, 5-13.
46. Arnold, E., Jacobo-Molina, A., Nanni, R. G., Williams, R. L., Lu, X., Ding, J., Clark, A. D., Jr., Zhang, A., Ferris, A. L., Clark, P., and et al. (1992) Structure of HIV-1 reverse transcriptase/DNA complex at 7 A resolution showing active site locations, *Nature 357*, 85-89.
47. Ding, J., Das, K., Hsiou, Y., Sarafianos, S. G., Clark, A. D., Jr., Jacobo-Molina, A., Tantillo, C., Hughes, S. H., and Arnold, E. (1998) Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8 A resolution, *J Mol Biol 284*, 1095-1111.
48. Xiong, Y., and Eickbush, T. H. (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences, *EMBO J 9*, 3353-3362.
49. Julias, J. G., McWilliams, M. J., Sarafianos, S. G., Arnold, E., and Hughes, S. H. (2002) Mutations in the RNase H domain of HIV-1 reverse transcriptase affect the initiation of DNA synthesis and the specificity of RNase H cleavage in vivo, *Proc Natl Acad Sci U S A 99*, 9515-9520.
50. Rausch, J. W., Lener, D., Miller, J. T., Julias, J. G., Hughes, S. H., and Le Grice, S. F. (2002) Altering the RNase H primer grip of human immunodeficiency virus reverse transcriptase modifies cleavage specificity, *Biochemistry 41*, 4856-4865.
51. Das, K., Martinez, S. E., Bandwar, R. P., and Arnold, E. (2014) Structures of HIV-1 RT-RNA/DNA ternary complexes with dATP and nevirapine reveal conformational flexibility of RNA/DNA: insights into requirements for RNase H cleavage, *Nucleic Acids Res 42*, 8125-8137.
52. Boyer, J. C., Bebenek, K., and Kunkel, T. A. (1992) Unequal human immunodeficiency virus type 1 reverse transcriptase error rates with RNA and DNA templates, *Proc Natl Acad Sci U S A 89*, 6919-6923.
53. Kerr, S. G., and Anderson, K. S. (1997) RNA dependent DNA replication fidelity of HIV-1 reverse transcriptase: evidence of discrimination between DNA and RNA substrates, *Biochemistry 36*, 14056-14063.
54. Lapkouski, M., Tian, L., Miller, J. T., Le Grice, S. F., and Yang, W. (2013) Complexes of HIV-1 RT, NNRTI and RNA/DNA hybrid reveal a structure compatible with RNA degradation, *Nat Struct Mol Biol 20*, 230-236.
55. Hsiou, Y., Ding, J., Das, K., Clark, A. D., Jr., Hughes, S. H., and Arnold, E. (1996) Structure of unliganded HIV-1 reverse transcriptase at 2.7 A resolution: implications of conformational changes for polymerization and inhibition mechanisms, *Structure 4*, 853-860.

56. Rodgers, D. W., Gamblin, S. J., Harris, B. A., Ray, S., Culp, J. S., Hellmig, B., Woolf, D. J., Debouck, C., and Harrison, S. C. (1995) The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1, *Proc Natl Acad Sci U S A 92*, 1222-1226.
57. Sarafianos, S. G., Das, K., Clark, A. D., Jr., Ding, J., Boyer, P. L., Hughes, S. H., and Arnold, E. (1999) Lamivudine (3TC) resistance in HIV-1 reverse transcriptase involves steric hindrance with beta-branched amino acids, *Proc Natl Acad Sci U S A 96*, 10027-10032.
58. Kirmizialtin, S., Nguyen, V., Johnson, K. A., and Elber, R. (2012) How conformational dynamics of DNA polymerase select correct substrates: experiments and simulations, *Structure 20*, 618-627.
59. Sarafianos, S. G., Clark, A. D., Jr., Das, K., Tuske, S., Birktoft, J. J., Ilankumaran, P., Ramesha, A. R., Sayer, J. M., Jerina, D. M., Boyer, P. L., Hughes, S. H., and Arnold, E. (2002) Structures of HIV-1 reverse transcriptase with pre- and post-translocation AZTMP-terminated DNA, *EMBO J 21*, 6614-6624.
60. Havlir, D. V., and Richman, D. D. (1996) Viral dynamics of HIV: implications for drug development and therapeutic strategies, *Ann Intern Med 124*, 984-994.
61. Mitsuya, H., Yarchoan, R., and Broder, S. (1990) Molecular targets for AIDS therapy, *Science 249*, 1533-1544.
62. Perno, C. F., Yarchoan, R., Cooney, D. A., Hartman, N. R., Gartner, S., Popovic, M., Hao, Z., Gerrard, T. L., Wilson, Y. A., Johns, D. G., and et al. (1988) Inhibition of human immunodeficiency virus (HIV-1/HTLV-IIIBa-L) replication in fresh and cultured human peripheral blood monocytes/macrophages by azidothymidine and related 2',3'-dideoxynucleosides, *J Exp Med 168*, 1111-1125.
63. Johnson, A. A., Ray, A. S., Hanes, J., Suo, Z., Colacino, J. M., Anderson, K. S., and Johnson, K. A. (2001) Toxicity of antiviral nucleoside analogs and the human mitochondrial DNA polymerase, *J Biol Chem 276*, 40847-40857.
64. Lee, H., Hanes, J., and Johnson, K. A. (2003) Toxicity of nucleoside analogues used to treat AIDS and the selectivity of the mitochondrial DNA polymerase, *Biochemistry 42*, 14711-14719.
65. Kewn, S., Hoggard, P. G., Henry-Mowatt, J. S., Veal, G. J., Sales, S. D., Barry, M. G., and Back, D. J. (1999) Intracellular activation of 2',3'-dideoxyinosine and drug interactions in vitro, *AIDS Res Hum Retroviruses 15*, 793-802.
66. Palella, F. J., Jr., Delaney, K. M., Moorman, A. C., Loveless, M. O., Fuhrer, J., Satten, G. A., Aschman, D. J., and Holmberg, S. D. (1998) Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators, *N Engl J Med 338*, 853-860.
67. Johnson, V. A., Brun-Vezinet, F., Clotet, B., Gunthard, H. F., Kuritzkes, D. R., Pillay, D., Schapiro, J. M., and Richman, D. D. (2010) Update of the drug resistance mutations in HIV-1: December 2010, *Top HIV Med 18*, 156-163.
68. Johnson, V. A., Calvez, V., Gunthard, H. F., Paredes, R., Pillay, D., Shafer, R., Wensing, A. M., and Richman, D. D. (2011) 2011 update of the drug resistance mutations in HIV-1, *Top Antivir Med 19*, 156-164.

69. Clavel, F., and Hance, A. J. (2004) HIV drug resistance, *N Engl J Med 350*, 1023-1035.
70. Das, K., Bandwar, R. P., White, K. L., Feng, J. Y., Sarafianos, S. G., Tuske, S., Tu, X., Clark, A. D., Jr., Boyer, P. L., Hou, X., Gaffney, B. L., Jones, R. A., Miller, M. D., Hughes, S. H., and Arnold, E. (2009) Structural basis for the role of the K65R mutation in HIV-1 reverse transcriptase polymerization, excision antagonism, and tenofovir resistance, *J Biol Chem 284*, 35092-35100.
71. Parikh, U. M., Bacheler, L., Koontz, D., and Mellors, J. W. (2006) The K65R mutation in human immunodeficiency virus type 1 reverse transcriptase exhibits bidirectional phenotypic antagonism with thymidine analog mutations, *J Virol 80*, 4971-4977.
72. Garforth, S. J., Domaoal, R. A., Lwatula, C., Landau, M. J., Meyer, A. J., Anderson, K. S., and Prasad, V. R. (2010) K65R and K65A substitutions in HIV-1 reverse transcriptase enhance polymerase fidelity by decreasing both dNTP misinsertion and mispaired primer extension efficiencies, *J Mol Biol 401*, 33-44.
73. Larder, B. A., Kemp, S. D., and Harrigan, P. R. (1995) Potential mechanism for sustained antiretroviral efficacy of AZT-3TC combination therapy, *Science 269*, 696-699.
74. Jérôme Courcambeck, R. P., Jérôme Deval, Boulbaba Selmi, Bruno Canard, Philippe Halfon, Michael D. Miller and Gérard Pèpe Molecular Mechanism of Resistance of Tenofovir against HIV RT K65R Strain and Mechanism of Resistance Suppression by M184V Mutation.
75. Gotte, M., Arion, D., Parniak, M. A., and Wainberg, M. A. (2000) The M184V mutation in the reverse transcriptase of human immunodeficiency virus type 1 impairs rescue of chain-terminated DNA synthesis, *J Virol 74*, 3579-3585.
76. Tu, X., Das, K., Han, Q., Bauman, J. D., Clark, A. D., Jr., Hou, X., Frenkel, Y. V., Gaffney, B. L., Jones, R. A., Boyer, P. L., Hughes, S. H., Sarafianos, S. G., and Arnold, E. (2010) Structural basis of HIV-1 resistance to AZT by excision, *Nat Struct Mol Biol 17*, 1202-1209.
77. Kellinger, M. W., and Johnson, K. A. (2011) Role of induced fit in limiting discrimination against AZT by HIV reverse transcriptase, *Biochemistry 50*, 5008-5015.
78. Ray, A. S., Murakami, E., Basavapathruni, A., Vaccaro, J. A., Ulrich, D., Chu, C. K., Schinazi, R. F., and Anderson, K. S. (2003) Probing the molecular mechanisms of AZT drug resistance mediated by HIV-1 reverse transcriptase using a transient kinetic analysis, *Biochemistry 42*, 8831-8841.
79. Boyer, P. L., Sarafianos, S. G., Arnold, E., and Hughes, S. H. (2001) Selective excision of AZTMP by drug-resistant human immunodeficiency virus reverse transcriptase, *J Virol 75*, 4832-4842.
80. Sluis-Cremer, N., Arion, D., Parikh, U., Koontz, D., Schinazi, R. F., Mellors, J. W., and Parniak, M. A. (2005) The 3'-azido group is not the primary determinant of 3'-azido-3'-deoxythymidine (AZT) responsible for the excision phenotype of AZT-resistant HIV-1, *J Biol Chem 280*, 29047-29052.

81. Spence, R. A., Kati, W. M., Anderson, K. S., and Johnson, K. A. (1995) Mechanism of inhibition of HIV-1 reverse transcriptase by nonnucleoside inhibitors, *Science 267*, 988-993.
82. Esnouf, R., Ren, J., Ross, C., Jones, Y., Stammers, D., and Stuart, D. (1995) Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors, *Nat Struct Biol 2*, 303-308.
83. Wang, J., Zhang, G., Bambara, R. A., Li, D., Liang, H., Wu, H., Smith, H. M., Lowe, N. R., Demeter, L. M., and Dykes, C. (2011) Nonnucleoside reverse transcriptase inhibitor-resistant HIV is stimulated by efavirenz during early stages of infection, *J Virol 85*, 10861-10873.
84. Johnson, K. A. (1993) Conformational coupling in DNA polymerase fidelity, *Annu Rev Biochem 62*, 685-713.
85. Johnson, K. A. (2010) The kinetic and chemical mechanism of high-fidelity DNA polymerases, *Biochim Biophys Acta 1804*, 1041-1048.
86. Kellinger, M. W., and Johnson, K. A. (2010) Nucleotide-dependent conformational change governs specificity and analog discrimination by HIV reverse transcriptase, *Proc Natl Acad Sci U S A 107*, 7734-7739.
87. Tsai, Y. C., and Johnson, K. A. (2006) A new paradigm for DNA polymerase specificity, *Biochemistry 45*, 9675-9687.
88. Patel, S. S., Wong, I., and Johnson, K. A. (1991) Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant, *Biochemistry 30*, 511-525.
89. Hanes, J. W., Thal, D. M., and Johnson, K. A. (2006) Incorporation and replication of 8-oxo-deoxyguanosine by the human mitochondrial DNA polymerase, *J Biol Chem 281*, 36241-36248.
90. Hanes, J. W., and Johnson, K. A. (2007) A novel mechanism of selectivity against AZT by the human mitochondrial DNA polymerase, *Nucleic Acids Res 35*, 6973-6983.
91. Batabyal, D., McKenzie, J. L., and Johnson, K. A. (2010) Role of histidine 932 of the human mitochondrial DNA polymerase in nucleotide discrimination and inherited disease, *J Biol Chem 285*, 34191-34201.
92. Gopalakrishnan, V., Peliska, J. A., and Benkovic, S. J. (1992) Human immunodeficiency virus type 1 reverse transcriptase: spatial and temporal relationship between the polymerase and RNase H activities, *Proc Natl Acad Sci U S A 89*, 10763-10767.
93. Beilhartz, G. L., Wendeler, M., Baichoo, N., Rausch, J., Le Grice, S., and Gotte, M. (2009) HIV-1 reverse transcriptase can simultaneously engage its DNA/RNA substrate at both DNA polymerase and RNase H active sites: implications for RNase H inhibition, *J Mol Biol 388*, 462-474.
94. Purohit, V., Roques, B. P., Kim, B., and Bambara, R. A. (2007) Mechanisms that prevent template inactivation by HIV-1 reverse transcriptase RNase H cleavages, *J Biol Chem 282*, 12598-12609.

95. Shaw-Reid, C. A., Feuston, B., Munshi, V., Getty, K., Krueger, J., Hazuda, D. J., Parniak, M. A., Miller, M. D., and Lewis, D. (2005) Dissecting the effects of DNA polymerase and ribonuclease H inhibitor combinations on HIV-1 reverse-transcriptase activities, *Biochemistry 44*, 1595-1606.
96. Abbondanzieri, E. A., Bokinsky, G., Rausch, J. W., Zhang, J. X., Le Grice, S. F., and Zhuang, X. (2008) Dynamic binding orientations direct activity of HIV reverse transcriptase, *Nature 453*, 184-189.
97. Liu, S., Abbondanzieri, E. A., Rausch, J. W., Le Grice, S. F., and Zhuang, X. (2008) Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates, *Science 322*, 1092-1097.
98. Suo, Z., and Johnson, K. A. (1997) Effect of RNA secondary structure on the kinetics of DNA synthesis catalyzed by HIV-1 reverse transcriptase, *Biochemistry 36*, 12459-12467.
99. Suo, Z., and Johnson, K. A. (1998) DNA secondary structure effects on DNA synthesis catalyzed by HIV-1 reverse transcriptase, *J Biol Chem 273*, 27259-27267.
100. Kim, S., Schroeder, C. M., and Xie, X. S. (2010) Single-molecule study of DNA polymerization activity of HIV-1 reverse transcriptase on DNA templates, *J Mol Biol 395*, 995-1006.
101. Suo, Z., and Johnson, K. A. (1997) Effect of RNA secondary structure on RNA cleavage catalyzed by HIV-1 reverse transcriptase, *Biochemistry 36*, 12468-12476.
102. Suo, Z., and Johnson, K. A. (1997) RNA secondary structure switching during DNA synthesis catalyzed by HIV-1 reverse transcriptase, *Biochemistry 36*, 14778-14785.
103. Isel, C., Lanchy, J. M., Le Grice, S. F., Ehresmann, C., Ehresmann, B., and Marquet, R. (1996) Specific initiation and switch to elongation of human immunodeficiency virus type 1 reverse transcription require the post-transcriptional modifications of primer tRNA3Lys, *EMBO J 15*, 917-924.
104. Lanchy, J. M., Keith, G., Le Grice, S. F., Ehresmann, B., Ehresmann, C., and Marquet, R. (1998) Contacts between reverse transcriptase and the primer strand govern the transition from initiation to elongation of HIV-1 reverse transcription, *J Biol Chem 273*, 24425-24432.
105. Liu, S., Harada, B. T., Miller, J. T., Le Grice, S. F., and Zhuang, X. (2010) Initiation complex dynamics direct the transitions between distinct phases of early HIV reverse transcription, *Nat Struct Mol Biol 17*, 1453-1460.
106. Le Grice, S. F. (1993) Human immunodeficiency virus reverse transcritptase, pp 163-191, Cold Spring Harbor laboratory Pess, NY.
107. Telesnitsky, A., and Goff, S. P. (1997) Reverse Transcriptase and the Generation of Retroviral DNA.
108. Larder, B. A., and Kemp, S. D. (1989) Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT), *Science 246*, 1155-1158.

109. Larder, B. A., Bloor, S., Kemp, S. D., Hertogs, K., Desmet, R. L., Miller, V., Sturmer, M., Staszewski, S., Ren, J., Stammers, D. K., Stuart, D. I., and Pauwels, R. (1999) A family of insertion mutations between codons 67 and 70 of human immunodeficiency virus type 1 reverse transcriptase confer multinucleoside analog resistance, *Antimicrob Agents Chemother 43*, 1961-1967.
110. De Antoni, A., Foli, A., Lisziewicz, J., and Lori, F. (1997) Mutations in the pol gene of human immunodeficiency virus type 1 in infected patients receiving didanosine and hydroxyurea combination therapy, *J Infect Dis 176*, 899-903.
111. de Jong, J. J., Goudsmit, J., Lukashov, V. V., Hillebrand, M. E., Baan, E., Huismans, R., Danner, S. A., ten Veen, J. H., de Wolf, F., and Jurriaans, S. (1999) Insertion of two amino acids combined with changes in reverse transcriptase containing tyrosine-215 of HIV-1 resistant to multiple nucleoside analogs, *AIDS 13*, 75-80.
112. Masquelier, B., Race, E., Tamalet, C., Descamps, D., Izopet, J., Buffet-Janvresse, C., Ruffault, A., Mohammed, A. S., Cottalorda, J., Schmuck, A., Calvez, V., Dam, E., Fleury, H., and Brun-Vezinet, F. (2001) Genotypic and phenotypic resistance patterns of human immunodeficiency virus type 1 variants with insertions or deletions in the reverse transcriptase (RT): multicenter study of patients treated with RT inhibitors, *Antimicrob Agents Chemother 45*, 1836-1842.
113. Tamalet, C., Yahi, N., Tourres, C., Colson, P., Quinson, A. M., Poizot-Martin, I., Dhiver, C., and Fantini, J. (2000) Multidrug resistance genotypes (insertions in the beta3-beta4 finger subdomain and MDR mutations) of HIV-1 reverse transcriptase from extensively treated patients: incidence and association with other resistance mutations, *Virology 270*, 310-316.
114. Menendez-Arias, L., Matamoros, T., and Cases-Gonzalez, C. E. (2006) Insertions and deletions in HIV-1 reverse transcriptase: consequences for drug resistance and viral fitness, *Curr Pharm Des 12*, 1811-1825.
115. Matamoros, T., Franco, S., Vazquez-Alvarez, B. M., Mas, A., Martinez, M. A., and Menendez-Arias, L. (2004) Molecular determinants of multi-nucleoside analogue resistance in HIV-1 reverse transcriptases containing a dipeptide insertion in the fingers subdomain: effect of mutations D67N and T215Y on removal of thymidine nucleotide analogues from blocked DNA primers, *J Biol Chem 279*, 24569-24577.
116. Cases-Gonzalez, C. E., Franco, S., Martinez, M. A., and Menendez-Arias, L. (2007) Mutational patterns associated with the 69 insertion complex in multi-drug-resistant HIV-1 reverse transcriptase that confer increased excision activity and high-level resistance to zidovudine, *J Mol Biol 365*, 298-309.
117. Boyer, P. L., Sarafianos, S. G., Arnold, E., and Hughes, S. H. (2002) Nucleoside analog resistance caused by insertions in the fingers of human immunodeficiency virus type 1 reverse transcriptase involves ATP-mediated excision, *J Virol 76*, 9143-9151.
118. Lennerstrand, J., Hertogs, K., Stammers, D. K., and Larder, B. A. (2001) Correlation between viral resistance to zidovudine and resistance at the reverse

transcriptase level for a panel of human immunodeficiency virus type 1 mutants, *J Virol 75*, 7202-7205.
119. Mas, A., Parera, M., Briones, C., Soriano, V., Martinez, M. A., Domingo, E., and Menendez-Arias, L. (2000) Role of a dipeptide insertion between codons 69 and 70 of HIV-1 reverse transcriptase in the mechanism of AZT resistance, *EMBO J 19*, 5752-5761.
120. Meyer, P. R., Lennerstrand, J., Matsuura, S. E., Larder, B. A., and Scott, W. A. (2003) Effects of dipeptide insertions between codons 69 and 70 of human immunodeficiency virus type 1 reverse transcriptase on primer unblocking, deoxynucleoside triphosphate inhibition, and DNA chain elongation, *J Virol 77*, 3871-3877.
121. Winters, M. A., Coolley, K. L., Girard, Y. A., Levee, D. J., Hamdan, H., Shafer, R. W., Katzenstein, D. A., and Merigan, T. C. (1998) A 6-basepair insert in the reverse transcriptase gene of human immunodeficiency virus type 1 confers resistance to multiple nucleoside inhibitors, *J Clin Invest 102*, 1769-1775.
122. Meyer, P. R., Matsuura, S. E., Mian, A. M., So, A. G., and Scott, W. A. (1999) A mechanism of AZT resistance: an increase in nucleotide-dependent primer unblocking by mutant HIV-1 reverse transcriptase, *Mol Cell 4*, 35-43.
123. Lennerstrand, J., Stammers, D. K., and Larder, B. A. (2001) Biochemical mechanism of human immunodeficiency virus type 1 reverse transcriptase resistance to stavudine, *Antimicrob Agents Chemother 45*, 2144-2146.
124. Mas, A., Vazquez-Alvarez, B. M., Domingo, E., and Menendez-Arias, L. (2002) Multidrug-resistant HIV-1 reverse transcriptase: involvement of ribonucleotide-dependent phosphorolysis in cross-resistance to nucleoside analogue inhibitors, *J Mol Biol 323*, 181-197.
125. Hanes, J. W., and Johnson, K. A. (2008) Real-time measurement of pyrophosphate release kinetics, *Anal Biochem 372*, 125-127.
126. Meyer, P. R., Matsuura, S. E., So, A. G., and Scott, W. A. (1998) Unblocking of chain-terminated primer by HIV-1 reverse transcriptase through a nucleotide-dependent mechanism, *Proc Natl Acad Sci U S A 95*, 13471-13476.
127. Tong, W., Lu, C. D., Sharma, S. K., Matsuura, S., So, A. G., and Scott, W. A. (1997) Nucleotide-induced stable complex formation by HIV-1 reverse transcriptase, *Biochemistry 36*, 5749-5757.
128. Curr, K., Tripathi, S., Lennerstrand, J., Larder, B. A., and Prasad, V. R. (2006) Influence of naturally occurring insertions in the fingers subdomain of human immunodeficiency virus type 1 reverse transcriptase on polymerase fidelity and mutation frequencies in vitro, *J Gen Virol 87*, 419-428.
129. Marchand, B., and Gotte, M. (2003) Site-specific footprinting reveals differences in the translocation status of HIV-1 reverse transcriptase. Implications for polymerase translocation and drug resistance, *J Biol Chem 278*, 35362-35372.
130. Gotte, M., Maier, G., Gross, H. J., and Heumann, H. (1998) Localization of the active site of HIV-1 reverse transcriptase-associated RNase H domain on a DNA

template using site-specific generated hydroxyl radicals, *J Biol Chem 273*, 10139-10146.
131. Traut, T. W. (1994) Physiological concentrations of purines and pyrimidines, *Mol Cell Biochem 140*, 1-22.
132. Meyer, P. R., Matsuura, S. E., Tolun, A. A., Pfeifer, I., So, A. G., Mellors, J. W., and Scott, W. A. (2002) Effects of specific zidovudine resistance mutations and substrate structure on nucleotide-dependent primer unblocking by human immunodeficiency virus type 1 reverse transcriptase, *Antimicrob Agents Chemother 46*, 1540-1545.
133. Nguyen, L. A., Daddacha, W., Rigby, S., Bambara, R. A., and Kim, B. (2012) Altered strand transfer activity of a multiple-drug-resistant human immunodeficiency virus type 1 reverse transcriptase mutant with a dipeptide fingers domain insertion, *J Mol Biol 415*, 248-262.
134. Le Grice, S. F. (1993) Human immunodeficiency virus reverse transcritptase, pp 163-191, Cold Spring Harbor laboratory Pess, NY.
135. Telesnitsky, A., and Goff, S. P. (1997) Reverse Transcriptase and the Generation of Retroviral DNA, In Retroviruses, J.M. Coffin, S.H.Hughes, and H. E. Varmus ed., pp 121-160, Cold Spring Harbor Press, NY.
136. World Health Organization. Scaling Up Antiretroviral Therapy in Renounce Limited Settings: Treatment Guidelines for a Public Health Approach -2003. URL: http://www.who.int/hiv/pub/prev_care/en/ScalingUp_E.pdf.
137. Rooke, R., Parniak, M. A., Tremblay, M., Soudeyns, H., Li, X. G., Gao, Q., Yao, X. J., and Wainberg, M. A. (1991) Biological comparison of wild-type and zidovudine-resistant isolates of human immunodeficiency virus type 1 from the same subjects: susceptibility and resistance to other drugs, *Antimicrob Agents Chemother 35*, 988-991.
138. Kellam, P., Boucher, C. A., and Larder, B. A. (1992) Fifth mutation in human immunodeficiency virus type 1 reverse transcriptase contributes to the development of high-level resistance to zidovudine, *Proc Natl Acad Sci U S A 89*, 1934-1938.
139. Kerr, S. G., and Anderson, K. S. (1997) Pre-steady-state kinetic characterization of wild type and 3'-azido-3'-deoxythymidine (AZT) resistant human immunodeficiency virus type 1 reverse transcriptase: implication of RNA directed DNA polymerization in the mechanism of AZT resistance, *Biochemistry 36*, 14064-14070.
140. Krebs, R., Immendorfer, U., Thrall, S. H., Wohrl, B. M., and Goody, R. S. (1997) Single-step kinetics of HIV-1 reverse transcriptase mutants responsible for virus resistance to nucleoside inhibitors zidovudine and 3-TC, *Biochemistry 36*, 10292-10300.
141. Meyer, P. R., Smith, A. J., Matsuura, S. E., and Scott, W. A. (2004) Effects of primer-template sequence on ATP-dependent removal of chain-terminating nucleotide analogues by HIV-1 reverse transcriptase, *J Biol Chem 279*, 45389-45398.

142. Victorova, L., Sosunov, V., Skoblov, A., Shipytsin, A., and Krayevsky, A. (1999) New substrates of DNA polymerases, *FEBS Lett 453*, 6-10.
143. Meyer, P. R., Smith, A. J., Matsuura, S. E., and Scott, W. A. (2006) Chain-terminating dinucleoside tetraphosphates are substrates for DNA polymerization by human immunodeficiency virus type 1 reverse transcriptase with increased activity against thymidine analogue-resistant mutants, *Antimicrob Agents Chemother 50*, 3607-3614.
144. Meyer, P. R., Matsuura, S. E., Schinazi, R. F., So, A. G., and Scott, W. A. (2000) Differential removal of thymidine nucleotide analogues from blocked DNA chains by human immunodeficiency virus reverse transcriptase in the presence of physiological concentrations of 2'-deoxynucleoside triphosphates, *Antimicrob Agents Chemother 44*, 3465-3472.
145. Meteer, J. D., Schinazi, R. F., Mellors, J. W., and Sluis-Cremer, N. (2014) Molecular mechanism of HIV-1 resistance to 3'-azido-2',3'-dideoxyguanosine, *Antiviral Res 101*, 62-67.
146. Meteer, J. D., Koontz, D., Asif, G., Zhang, H. W., Detorio, M., Solomon, S., Coats, S. J., Sluis-Cremer, N., Schinazi, R. F., and Mellors, J. W. (2011) The base component of 3'-azido-2',3'-dideoxynucleosides influences resistance mutations selected in HIV-1 reverse transcriptase, *Antimicrob Agents Chemother 55*, 3758-3764.
147. Hostomsky, Z., Hostomska, Z., Hudson, G. O., Moomaw, E. W., and Nodes, B. R. (1991) Reconstitution in vitro of RNase H activity by using purified N-terminal and C-terminal domains of human immunodeficiency virus type 1 reverse transcriptase, *Proc Natl Acad Sci U S A 88*, 1148-1152.
148. Evans, D. B., Brawn, K., Deibel, M. R., Jr., Tarpley, W. G., and Sharma, S. K. (1991) A recombinant ribonuclease H domain of HIV-1 reverse transcriptase that is enzymatically active, *J Biol Chem 266*, 20583-20585.
149. Stammers, D. K., Tisdale, M., Court, S., Parmar, V., Bradley, C., and Ross, C. K. (1991) Rapid purification and characterisation of HIV-1 reverse transcriptase and RNaseH engineered to incorporate a C-terminal tripeptide alpha-tubulin epitope, *FEBS Lett 283*, 298-302.
150. Smith, J. S., and Roth, M. J. (1993) Purification and characterization of an active human immunodeficiency virus type 1 RNase H domain, *J Virol 67*, 4037-4049.
151. Stahl, S. J., Kaufman, J. D., Vikic-Topic, S., Crouch, R. J., and Wingfield, P. T. (1994) Construction of an enzymatically active ribonuclease H domain of human immunodeficiency virus type 1 reverse transcriptase, *Protein Eng 7*, 1103-1108.
152. Keck, J. L., and Marqusee, S. (1995) Substitution of a highly basic helix/loop sequence into the RNase H domain of human immunodeficiency virus reverse transcriptase restores its Mn(2+)-dependent RNase H activity, *Proc Natl Acad Sci U S A 92*, 2740-2744.
153. Nowotny, M., Gaidamakov, S. A., Ghirlando, R., Cerritelli, S. M., Crouch, R. J., and Yang, W. (2007) Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription, *Mol Cell 28*, 264-276.

154. Auerbach, S. D., and Johnson, K. A. (2005) Alternating site ATPase pathway of rat conventional kinesin, *J Biol Chem 280*, 37048-37060.
155. Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data, *Anal Biochem 387*, 30-41.
156. Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) Global kinetic explorer: a new computer program for dynamic simulation and fitting of kinetic data, *Anal Biochem 387*, 20-29.
157. Schultz, S. J., Zhang, M., and Champoux, J. J. (2010) Multiple nucleotide preferences determine cleavage-site recognition by the HIV-1 and M-MuLV RNases H, *J Mol Biol 397*, 161-178.
158. Marchand, B., Tchesnokov, E. P., and Gotte, M. (2007) The pyrophosphate analogue foscarnet traps the pre-translocational state of HIV-1 reverse transcriptase in a Brownian ratchet model of polymerase translocation, *J Biol Chem 282*, 3337-3346.
159. DeStefano, J. J., Buiser, R. G., Mallaber, L. M., Bambara, R. A., and Fay, P. J. (1991) Human immunodeficiency virus reverse transcriptase displays a partially processive 3' to 5' endonuclease activity, *J Biol Chem 266*, 24295-24301.
160. Krug, M. S., and Berger, S. L. (1989) Ribonuclease H activities associated with viral reverse transcriptases are endonucleases, *Proc Natl Acad Sci U S A 86*, 3539-3543.
161. Beilhartz, G. L., and Gotte, M. (2010) HIV-1 Ribonuclease H: Structure, Catalytic Mechanism and Inhibitors, *Viruses 2*, 900-926.
162. Furfine, E. S., and Reardon, J. E. (1991) Reverse transcriptase.RNase H from the human immunodeficiency virus. Relationship of the DNA polymerase and RNA hydrolysis activities, *J Biol Chem 266*, 406-412.
163. Champoux, J. J., and Schultz, S. J. (2009) Ribonuclease H: properties, substrate specificity and roles in retroviral reverse transcription, *FEBS J 276*, 1506-1516.
164. Wohrl, B. M., Krebs, R., Goody, R. S., and Restle, T. (1999) Refined model for primer/template binding by HIV-1 reverse transcriptase: pre-steady-state kinetic analyses of primer/template binding and nucleotide incorporation events distinguish between different binding modes depending on the nature of the nucleic acid substrate, *J Mol Biol 292*, 333-344.
165. Gao, H. Q., Boyer, P. L., Arnold, E., and Hughes, S. H. (1998) Effects of mutations in the polymerase domain on the polymerase, RNase H and strand transfer activities of human immunodeficiency virus type 1 reverse transcriptase, *J Mol Biol 277*, 559-572.
166. Powell, M. D., Beard, W. A., Bebenek, K., Howard, K. J., Le Grice, S. F., Darden, T. A., Kunkel, T. A., Wilson, S. H., and Levin, J. G. (1999) Residues in the alphaH and alphaI helices of the HIV-1 reverse transcriptase thumb subdomain required for the specificity of RNase H-catalyzed removal of the polypurine tract primer, *J Biol Chem 274*, 19885-19893.

167. Mandal, D., Dash, C., Le Grice, S. F., and Prasad, V. R. (2006) Analysis of HIV-1 replication block due to substitutions at F61 residue of reverse transcriptase reveals additional defects involving the RNase H function, *Nucleic Acids Res 34*, 2853-2863.
168. Nowak, M. (1990) HIV mutation rate, *Nature 347*, 522.
169. Preston, B. D., Poiesz, B. J., and Loeb, L. A. (1988) Fidelity of HIV-1 reverse transcriptase, *Science 242*, 1168-1171.
170. Kantor, R., Shafer, R. W., Follansbee, S., Taylor, J., Shilane, D., Hurley, L., Nguyen, D. P., Katzenstein, D., and Fessel, W. J. (2004) Evolution of resistance to drugs in HIV-1-infected patients failing antiretroviral therapy, *AIDS 18*, 1503-1511.
171. Boutabout, M., Wilhelm, M., and Wilhelm, F. X. (2001) DNA synthesis fidelity by the reverse transcriptase of the yeast retrotransposon Ty1, *Nucleic Acids Res 29*, 2217-2222.
172. Avidan, O., Meer, M. E., Oz, I., and Hizi, A. (2002) The processivity and fidelity of DNA synthesis exhibited by the reverse transcriptase of bovine leukemia virus, *Eur J Biochem 269*, 859-867.
173. Bakhanashvili, M., and Hizi, A. (1993) Fidelity of DNA synthesis exhibited in vitro by the reverse transcriptase of the lentivirus equine infectious anemia virus, *Biochemistry 32*, 7559-7567.
174. Roberts, J. D., Preston, B. D., Johnston, L. A., Soni, A., Loeb, L. A., and Kunkel, T. A. (1989) Fidelity of two retroviral reverse transcriptases during DNA-dependent DNA synthesis in vitro, *Mol Cell Biol 9*, 469-476.
175. Roberts, J. D., Bebenek, K., and Kunkel, T. A. (1988) The accuracy of reverse transcriptase from HIV-1, *Science 242*, 1171-1173.
176. Rezende, L. F., Drosopoulos, W. C., and Prasad, V. R. (1998) The influence of 3TC resistance mutation M184I on the fidelity and error specificity of human immunodeficiency virus type 1 reverse transcriptase, *Nucleic Acids Res 26*, 3066-3072.
177. Boyer, P. L., and Hughes, S. H. (2000) Effects of amino acid substitutions at position 115 on the fidelity of human immunodeficiency virus type 1 reverse transcriptase, *J Virol 74*, 6494-6500.
178. Ji, J. P., and Loeb, L. A. (1992) Fidelity of HIV-1 reverse transcriptase copying RNA in vitro, *Biochemistry 31*, 954-958.
179. Ji, J., and Loeb, L. A. (1994) Fidelity of HIV-1 reverse transcriptase copying a hypervariable region of the HIV-1 env gene, *Virology 199*, 323-330.
180. Achuthan, V., Keith, B. J., Connolly, B. A., and DeStefano, J. J. (2014) Human immunodeficiency virus reverse transcriptase displays dramatically higher fidelity under physiological magnesium conditions in vitro, *J Virol 88*, 8514-8527.
181. Brune, M., Hunter, J. L., Corrie, J. E., and Webb, M. R. (1994) Direct, real-time measurement of rapid inorganic phosphate release using a novel fluorescent probe and its application to actomyosin subfragment 1 ATPase, *Biochemistry 33*, 8262-8271.

182. Tantillo, C., Ding, J., Jacobo-Molina, A., Nanni, R. G., Boyer, P. L., Hughes, S. H., Pauwels, R., Andries, K., Janssen, P. A., and Arnold, E. (1994) Locations of anti-AIDS drug binding sites and resistance mutations in the three-dimensional structure of HIV-1 reverse transcriptase. Implications for mechanisms of drug inhibition and resistance, *J Mol Biol 243*, 369-387.
183. Sarafianos, S. G., Pandey, V. N., Kaushik, N., and Modak, M. J. (1995) Glutamine 151 participates in the substrate dNTP binding function of HIV-1 reverse transcriptase, *Biochemistry 34*, 7207-7216.
184. Kaushik, N., Harris, D., Rege, N., Modak, M. J., Yadav, P. N., and Pandey, V. N. (1997) Role of glutamine-151 of human immunodeficiency virus type-1 reverse transcriptase in RNA-directed DNA synthesis, *Biochemistry 36*, 14430-14438.
185. Kaushik, N., Rege, N., Yadav, P. N., Sarafianos, S. G., Modak, M. J., and Pandey, V. N. (1996) Biochemical analysis of catalytically crucial aspartate mutants of human immunodeficiency virus type 1 reverse transcriptase, *Biochemistry 35*, 11536-11546.
186. Chunduri, H., Crumpacker, C., and Sharma, P. L. (2011) Reverse transcriptase mutation K65N confers a decreased replication capacity to HIV-1 in comparison to K65R due to a decreased RT processivity, *Virology 414*, 34-41.
187. Shirasaka, T., Kavlick, M. F., Ueno, T., Gao, W. Y., Kojima, E., Alcaide, M. L., Chokekijchai, S., Roy, B. M., Arnold, E., Yarchoan, R., and et al. (1995) Emergence of human immunodeficiency virus type 1 variants with resistance to multiple dideoxynucleosides in patients receiving therapy with dideoxynucleosides, *Proc Natl Acad Sci U S A 92*, 2398-2402.
188. Ueno, T., and Mitsuya, H. (1997) Comparative enzymatic study of HIV-1 reverse transcriptase resistant to 2',3'-dideoxynucleotide analogs using the single-nucleotide incorporation assay, *Biochemistry 36*, 1092-1099.
189. Ueno, T., Shirasaka, T., and Mitsuya, H. (1995) Enzymatic characterization of human immunodeficiency virus type 1 reverse transcriptase resistant to multiple 2',3'-dideoxynucleoside 5'-triphosphates, *J Biol Chem 270*, 23605-23611.
190. Pandey, V. N., Kaushik, N., Rege, N., Sarafianos, S. G., Yadav, P. N., and Modak, M. J. (1996) Role of methionine 184 of human immunodeficiency virus type-1 reverse transcriptase in the polymerase function and fidelity of DNA synthesis, *Biochemistry 35*, 2168-2179.
191. Wainberg, M. A., Drosopoulos, W. C., Salomon, H., Hsu, M., Borkow, G., Parniak, M., Gu, Z., Song, Q., Manne, J., Islam, S., Castriota, G., and Prasad, V. R. (1996) Enhanced fidelity of 3TC-selected mutant HIV-1 reverse transcriptase, *Science 271*, 1282-1285.
192. Gu, Z., Gao, Q., Li, X., Parniak, M. A., and Wainberg, M. A. (1992) Novel mutation in the human immunodeficiency virus type 1 reverse transcriptase gene that encodes cross-resistance to 2',3'-dideoxyinosine and 2',3'-dideoxycytidine, *J Virol 66*, 7128-7135.
193. Schinazi, R. F., Lloyd, R. M., Jr., Nguyen, M. H., Cannon, D. L., McMillan, A., Ilksoy, N., Chu, C. K., Liotta, D. C., Bazmi, H. Z., and Mellors, J. W. (1993)

Characterization of human immunodeficiency viruses resistant to oxathiolane-cytosine nucleosides, *Antimicrob Agents Chemother 37*, 875-881.
194. Tisdale, M., Kemp, S. D., Parry, N. R., and Larder, B. A. (1993) Rapid in vitro selection of human immunodeficiency virus type 1 resistant to 3'-thiacytidine inhibitors due to a mutation in the YMDD region of reverse transcriptase, *Proc Natl Acad Sci U S A 90*, 5653-5656.
195. Larder, B. A., Coates, K. E., and Kemp, S. D. (1991) Zidovudine-resistant human immunodeficiency virus selected by passage in cell culture, *J Virol 65*, 5232-5236.
196. Lewis, D. A., Bebenek, K., Beard, W. A., Wilson, S. H., and Kunkel, T. A. (1999) Uniquely altered DNA replication fidelity conferred by an amino acid change in the nucleotide binding pocket of human immunodeficiency virus type 1 reverse transcriptase, *J Biol Chem 274*, 32924-32930.
197. Wu, E. Y., and Beese, L. S. (2011) The structure of a high fidelity DNA polymerase bound to a mismatched nucleotide reveals an "ajar" intermediate conformation in the nucleotide selection mechanism, *J Biol Chem 286*, 19758-19767.
198. Krahn, J. M., Beard, W. A., and Wilson, S. H. (2004) Structural insights into DNA polymerase beta deterrents for misincorporation support an induced-fit mechanism for fidelity, *Structure 12*, 1823-1832.